

# Optimal Policies for Risk-Averse Electric Vehicle Charging with Spot Purchases

Daniel R. Jiang\* and Warren B. Powell\*

May 11, 2016

## Abstract

We consider the sequential decision problem faced by the manager of an electric vehicle (EV) charging station, who aims to satisfy the charging demand of the customer while minimizing cost. Since the total time needed to charge the EV up to capacity is typically *less* than the amount of time that the customer is away, there are opportunities to exploit electricity spot price variations within some time window. However, it is also true that the return time of the customer is uncertain, so there exists the risk of an insufficient charge. We formulate the problem as a finite horizon Markov decision process (MDP) and consider a risk-averse objective function by optimizing under a dynamic risk measure constructed using a convex combination of expected value and conditional value at risk (CVaR). For the first time in the literature, we provide an analysis of the effect that risk parameters, e.g., the risk-level  $\alpha$  used in CVaR, have on the structure of the optimal policy. We show that becoming more risk-averse in the dynamic risk measure sense corresponds to the intuitively appealing notion of becoming more risk-averse in the order thresholds of the optimal policy. This result allows us to develop computational techniques for approximating a *spectrum of risk-averse policies* generated by varying the parameters of the risk measure. Finally, numerical results for a case study using spot price data from California ISO (CAISO) are shown, where the Pareto optimality of our policies when measured against practical metrics of risk and reward is examined.

## 1 Introduction

The recent popularity of electric vehicles (EVs) has spurred significant research attention in the area of EV storage management from a number of perspectives. One stream of literature, e.g., [Roe et al. \[2009\]](#), [Clement-Nyys et al. \[2010\]](#), [Sundstrom and Binding \[2010\]](#), [Sortomme et al. \[2011\]](#), [Rotering and Ilic \[2011\]](#), [Gan et al. \[2013\]](#), [Li et al. \[2014\]](#), [Wei et al. \[2014\]](#), takes an aggregate view of EVs and provides possible solutions to the negative impacts of an overloaded electrical infrastructure. Broadly speaking, the literature considers both coordinated and decentralized scheduling approaches to accomplish *peak-shifting*, i.e., flattening the load profile due to EV charging in order to reduce stress on the grid.

A complementary line of research deals with EV charging policies at the level of individual charging stations, focusing on the optimization of metrics such as charging cost and

---

\*Department of Operations Research and Financial Engineering, Princeton University

quality of service (minimal blocking/waiting times in queue-based models). For example, [Zhang et al. \[2014\]](#) integrates a queueing approach with Markov decision process (MDP) theory to optimize mean waiting times under a cost constraint. In [Bayram et al. \[2011\]](#), [Koutsopoulos et al. \[2011\]](#), and [Karbasioun et al. \[2014\]](#), the authors propose optimal operating strategies for the setting where a charging station is paired with storage. The papers [Bashash et al. \[2010\]](#) and [Bashash et al. \[2011\]](#) use a detailed model of a lithium ion battery in order to optimize under the objective of cost of electricity and battery health degradation. The uncertainty in the use of EVs, i.e., the driving pattern of the customer, is taken into account in the charge management policy proposed by [Iversen et al. \[2014\]](#).

The motivation behind the problem examined in this paper is new, but most closely related to the second stream of literature dealing with individual charging stations. We consider the point of view of a firm that owns charging stations located in public areas, such as parking garages, gas stations, or hotels. A typical scenario is the following: a customer drops off an EV at a charging station and returns after an uncertain amount of time, with the expectation that the vehicle is adequately charged (depending on the power available from the charger and the size of the battery, the time may vary greatly). In a recently filed patent application by Google, this exact problem is described as follows:

Modern EVSEs are equipped with technology that allows them to communicate with electric vehicles to control the charging rate. This capability allows the charging station to communicate the amount of charging capacity available to the EV... However, in many cases it is not necessary for the EV to charge the battery at the full rate. For example, EV drivers may drive their car[s] to work and leave [them] plugged in all day, even though charging completes in only a few hours. [\[Wytock et al., 2016\]](#)

The decision and risk analysis problem here is to balance the risk of the customer returning to a undercharged vehicle and the reward of completing a satisfactory charge at a reasonable cost before the customer returns. Why is this an important problem to study? There are three main reasons, outlined below.

- In order to reap the environmental and societal benefits of EVs, we must overcome the main hurdle to widespread EV adoption, *range anxiety* — the fear of being stranded away from a charging station with an empty battery [Eberle and von Helmolt \[2010\]](#). To do so, we would require a dense network of charging stations located nationwide, but such infrastructure is not possible without an economic incentive. Unfortunately, studies, e.g., [Schroeder and Traber \[2012\]](#) and [Chang et al. \[2012\]](#), show that in the current climate, there is little incentive in owning and operating EV infrastructure. Therefore, it is crucial for us to study methods to make ownership of EV charging stations a more viable business model.
- There are several levels of charging rates: Level 1 charging provides 2 to 5 miles per

hour of charge; Level 2 charging provides 10 to 20 miles per hour of charge; and finally, DC fast charging provides 180 to 240 miles per hour of charge [U.S. Department of Energy, 2012]. As high powered chargers are becoming more widespread, it is now increasingly common for the customer to be away for a length of time that is *longer* than the amount of time needed to fully charge the EV. Given the high costs of the DC fast charging technology, estimated at \$65,000–\$70,000 compared to \$15,000–\$18,000 for a Level 2 charger [U.S. Department of Energy, 2012] (see also Bay Area Council [2011] and Schroeder and Traber [2012] for similar estimates), it is clear that optimal control policies can help ease the economics of such an investment. Indeed, our conversations with industry colleagues confirm that this is a real problem faced owners of EV charging stations today.

- As we previously discussed, in the event of widespread adoption of EVs, the resulting high demand for energy can negatively impact the electrical grid. In order to accurately assess the effect of EVs on the grid at large, it is important to first understand the behaviors of firms controlling the individual charging stations.

As discussed in Bansal [2015], there exist three primary ways that charging fees are assessed for consumers: time-based (e.g., Blink, The Electric Circuit, ShorePower), energy-based (e.g., Blink), and subscription or membership-based (e.g., NRG EVgo), with the first two options being the most prevalent. In this paper, we model a time-based access fee because it is the most appropriate given the questions that we are considering. Unrelated to our decision problem, generally speaking, time-based fees come with the following advantages: (1) customers have the incentive to vacate the charging unit rather than block it without penalty; (2) they are legal to implement without a license (unlike an energy-based fee); (3) they are easy to comprehend for consumers and there is no issue of measurement inaccuracy. For a detailed analysis, see U.S. Department of Energy [2013]. Incidentally, (1) makes a risk-based analysis of this problem particularly fitting.

We suppose that the firm can directly interact with a spot energy market. Our optimal policy would depend on the spot prices of electricity  $P_t$  in the following way: if  $P_t$  is high, then we have a possible incentive to delay charging and buy later at a lower price. Of course, the trade-off here is that whenever charging is delayed, the risk of a dissatisfied customer (one who returns to a partially charged vehicle) increases. The cost of a *dissatisfied customer* can be substantial due to lost future sales. A precise characterization of this cost is outside the scope of this paper, but as discussed in Hogan et al. [2003], it can depend on a number of factors including social effects and market share. In this paper, we employ a simple model of a *terminal cost*, denoted  $D_{t+1}$ , that encompasses the cost of lost future sales. Therefore, the three sources of uncertainty in our problem are the spot prices, the arrival time of the customer, and the terminal cost.

We propose using the framework of risk-averse MDPs with dynamic risk measures, introduced in Ruszczyński [2010], to capture the important issue of sequential risk in this

problem; such a model has not yet been considered in the EV literature. The beauty of this framework is that when a certain *time-consistency* property holds (see Cheridito et al. [2006] and Ruszczyński [2010]), a familiar Bellman recursion applies for the value functions of the optimal risk-averse policy, allowing us to apply similar solution techniques as in the risk-neutral case. In applications, such as Philpott and de Matos [2012], Philpott et al. [2013], Shapiro et al. [2013], Kozmík and Morton [2014], and Rudloff et al. [2014], a commonly employed dynamic risk measure is constructed by nesting (or composing) one-step risk measures that are a convex combination of the expected value and conditional value at risk, which is also the risk measure that we choose to study in this paper. The main contributions of this paper are summarized below.

- We present a novel risk-averse, sequential model of a problem in EV charging that exhibits an uncertain horizon. The model is formulated using the framework of risk-averse MDPs with dynamic risk measures, constructed using one step risk measures that are a convex combination of expectation and conditional value at risk (CVaR).
- We introduce the notion of a *spectrum of risk-averse policies*, which represent the optimal policies generated by varying the parameters of the risk measure. We argue if this set of policies is easily characterized, then a manager has the capability of choosing a theoretically justified policy (i.e., from this set) that satisfies the risk and capital requirements of the firm. This is in contrast to the common practice of heuristically tuning an optimal risk-neutral policy to obtain risk-aversion.
- The main result of the paper, Theorem 3, states that the charging thresholds of the optimal policy relax as we increase the parameters of the dynamic risk measure. In addition, we find that the charging threshold is nonincreasing in the current spot price and that similar structural results hold for other practical metrics of risk and reward (see Corollary 4). An analysis of the effect of the risk parameters has not yet been considered in the literature; we believe that such analyses increase the practical applicability of MDPs specified using dynamic risk measures.
- We make use of Theorem 3 and offer a method of approximating the spectrum of risk-averse policies using *polynomial optimization and sum of squares constraints* [Ahmadi and Majumdar, 2015]. The idea is that by solving a small number of risk-averse MDPs, Theorem 3 allows us to reliably approximate an entire set of risk-averse optimal policies that are easily accessible for practitioners.

The paper is organized as follows. First, we present in Section 2 the mathematical model of the problem. In Section 3, we provide a structural analysis of the model showing that the optimal policy is of the *order-up-to* type with a threshold that depends on the spot price and risk parameters. Next, in Section 4, we analyze the behavior of the threshold of the order-up-to policy as a function of the risk parameter vector  $\beta_t = (\lambda_t, \alpha_t)$  and the

spot price. Finally, in Section 5, we provide numerical work to illustrate our results for an instance of the EV charging problem on the California ISO (CAISO) spot market.

## 2 Mathematical Formulation

In this section, we first give a brief overview of dynamic risk measures in MDPs and then proceed to give a mathematical formulation of the EV problem.

### 2.1 Dynamic Risk Measures in Markov Decision Processes

The main idea of a risk-averse MDP is that the objective is a time-consistent dynamic risk measure  $\rho_{0,T}$  of the *sequence* of downstream costs  $C_1, C_2, \dots, C_T$ , resulting from decisions made at  $t = 0, 1, \dots, T-1$ . Let  $(\Omega, \mathcal{F}, \mathbf{P})$  be the probability space on which  $C_1, C_2, \dots, C_T$  are defined, where  $\{C_t\}$  is adapted to a filtration  $\{\mathcal{F}_t\}_{t=0}^T$ , with  $\{\emptyset, \Omega\} = \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots \subseteq \mathcal{F}_T \subseteq \mathcal{F}$ . Let  $\mathcal{Z}_t$  be the space of  $\mathcal{F}_t$  measurable random variables. The objective is written using the notation  $\rho_{0,T-1}(C_1, C_2, \dots, C_T)$ ; in the risk-neutral case, we would have  $\rho_{0,T-1}(C_1, C_2, \dots, C_T) = \mathbf{E}\left[\sum_{t=1}^T C_t\right]$ , but in general, we cannot assume such an additive form. The notion of time-consistency can be summarized as: if we prefer one stream of costs at some time in the future, and the costs from now until then are identical, then we must also prefer the same stream of costs *today*; see Ruszczyński [2010] for a precise definition. It can be shown that under some mild assumptions that time-consistency allows us to write the following nested formulation

$$\rho_{0,T-1}(C_1, C_2, \dots, C_T) = \rho_0(C_1 + \rho_1(C_2 + \dots + \rho_{T-1}(C_T) \dots)), \quad (1)$$

for some *one-step conditional risk measures*  $\rho_t : \mathcal{Z}_{t+1} \rightarrow \mathcal{Z}_t$ .

Now we move to a sequential decision setting and consider some fixed operating policy  $\pi$ . Let the cost sequence under  $\pi$  be given by  $C_0^\pi, C_1^\pi, \dots, C_T^\pi$ . The traditional risk-neutral optimization problem (see, e.g., Puterman [2014]) is

$$\min_{\pi \in \Pi} \mathbf{E} \left[ \sum_{t=1}^T C_t^\pi \right] = \min_{\pi \in \Pi} \mathbf{E}_0(C_1^\pi + \dots + \mathbf{E}_{T-1}(C_T^\pi) \dots), \quad (2)$$

where  $\mathbf{E}_t$  is shorthand for  $\mathbf{E}(\cdot | \mathcal{F}_t)$ . We define the risk-averse optimization problem as

$$\min_{\pi \in \Pi} \rho_{0,T-1}(C_1^\pi, C_2^\pi, \dots, C_T^\pi), \quad (3)$$

where the expectation  $\mathbf{E}$  is replaced with a dynamic risk measure  $\rho_{0,T-1}$ . Applying the nested formulation of (1), we see that the risk-averse formulation (3) is a clear analog of the risk-neutral case (2) with the conditional expectations replaced with one-step risk conditional risk measures:

$$\min_{\pi \in \Pi} \rho_0(C_1^\pi + \rho_1(C_2^\pi + \dots + \rho_{T-1}(C_T^\pi) \dots)). \quad (4)$$

In this paper, we take  $\rho_t$  to be the mean-CVaR risk measure, which we define now. Let  $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_{T-1})$  and  $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_{T-1})$  with  $\lambda_t \in [0, 1]$ ,  $\alpha_t \in (0, 1)$ , and  $\beta_t = (\lambda_t, \alpha_t)$ . We call  $\beta = \{\beta_t\}_{t=0}^{T-1}$  the *risk parameters*. The parameter  $\lambda_t$  is the weight on the tail risk (higher  $\lambda$  means more risk aversion) and  $\alpha_t$  controls the length of the tail to consider (higher  $\alpha_t$  means more risk aversion). The one-step risk measures take the form

$$\rho_{t,\beta}(X) = (1 - \lambda_t) \mathbf{E}_t[X] + \lambda_t \text{CVaR}_{t,\alpha}(X), \quad (5)$$

where  $X$  is a random cost in  $\mathcal{Z}_{t+1}$ ,  $\mathbf{E}_t[\cdot]$  is a conditional expectation given information up to time  $t$ , and  $\text{CVaR}_{t,\alpha}(\cdot)$  is conditional value at risk given information up to time  $t$ . It is given by (see [Rockafellar and Uryasev \[2000\]](#) and [Ruszczyński \[2010\]](#)).

$$\text{CVaR}_{t,\alpha}(X) = \inf_u \left\{ u + \frac{1}{1 - \alpha_t} \mathbf{E}_t[(X - u)^+] \right\}. \quad (6)$$

Thus, we henceforth consider the optimization problem (4) with  $\rho_t := \rho_{t,\beta}$  for all  $t$ .

## 2.2 Risk-Averse EV Charging Model

The horizon of the problem spans  $t = 0$  (when the EV first arrives) to  $t = T$ , with decisions potentially being made at  $t \in \{0, 1, \dots, T-1\}$ . We say “potentially” because the decision problem ends at an uncertain point in time, represented by a random variable  $\tau \in \{1, 2, \dots, T\}$ , when the customer returns to the vehicle. The time-based access fee to the charging station at time  $t$  is denoted by  $f_t$ . Suppose that for every time  $t$ , our decision is to choose a charging rate  $x \in [0, x_{\max}]$ . Assuming normalized units, we obtain  $x$  units of energy at each time period.

Now let us specify the dynamics of the spot price process  $\{P_t\}$ , which we assume is independent from  $\tau$ . The spot price  $P_t$  follows the mean-reverting *Ornstein–Uhlenbeck (OU)* process,  $dP_t = \kappa_P (\theta_P - P_t) dt + \sigma_P dW_t$ , where  $\kappa_P > 0$  is the mean reversion parameter,  $\theta_P$  is the long term mean of the process,  $\sigma_P > 0$  is the volatility, and  $W_t$  is a standard Brownian motion. It is well known that the conditional distribution of  $P_{t+1}$  given  $P_t = p$  is normal with mean  $\mu_P(p)$  and variance  $v_P$ , where

$$\mu_P(p) = \theta_P + (p - \theta_P) e^{-\kappa_P} \quad \text{and} \quad v_P = \frac{\sigma_P^2}{2\kappa_P} (1 - e^{-2\kappa_P}). \quad (7)$$

The OU model allows spot prices to become negative, a phenomenon that is known to occur with electricity prices (see, e.g., [Zhou et al. \[2015\]](#)). Negative prices, however, are not a major issue for the purposes of this paper. In fact, the results obtained in this paper can hold for other models of the spot price process as well. The crucial property needed is that  $P_{t+1} | P_t = p$  dominates  $P_{t+1} | P_t = p'$  in the first order (see [Shaked and Shanthikumar \[2007\]](#)) whenever  $p$  is larger than  $p'$ .

Let  $R_t \in [0, R_{\max}] =: \mathcal{R}$  be the amount of energy in storage at time  $t$ . The *transition*

function for the resource state variable is simply given by the equation

$$R_{t+1} = (R_t + x_t) \mathbf{1}_{\{t < \tau\}} + R_t \mathbf{1}_{\{t \geq \tau\}}, \quad (8)$$

provided that  $x_t$  is chosen in the constraint set  $\mathcal{X}(R_t) = \{x : 0 \leq x \leq \min\{R_{\max} - R_t, x_{\max}\}\}$ . Our goal is to find a policy that balances between making  $R_\tau$  close to  $R_{\max}$  while also incurring minimal cost.

As previously mentioned, we model the *future cost* of a (possibly dissatisfied) customer at time  $\tau$  as a normal distribution whose mean depends on  $R_\tau$  and  $P_\tau$ . Specifically, let us define the random variables  $D_{t+1}(r, p) \sim \mathcal{N}(m_t(r, p), \sigma_t(r, p)^2)$ . In practice, the distribution should mostly be sensitive to changes in  $r$ , but we allow dependence on  $p$  as well in order to capture the future evolution of spot prices. For instance, a high price today may signal higher prices in the near future. Therefore,  $D_{t+1}(r, p)$  is assumed to model both information about the lost future sales *and* the spot market.

Notice that we must only distinguish between the events  $\{t < \tau\}$ ,  $\{t = \tau\}$ , and  $\{t > \tau\}$ , so it is convenient for us to consider the auxiliary process  $A_t = \mathbf{1}_{\{t < \tau\}} - \mathbf{1}_{\{t > \tau\}}$  where  $A_t = 1$  means “the customer has not yet returned,”  $A_t = 0$  means “customer is currently at the station and will leave in the next time period,” and  $A_t = -1$  means “the customer has taken the EV and left.” The dynamics of this process are as follows. First, it is clear that several transitions occur with probability one: if  $A_t = 0$ , then  $A_{t+1} = -1$  and similarly, if  $A_t = -1$ , then  $A_{t+1} = -1$ . The remaining case is to find  $q_t^0 := \mathbf{P}(A_{t+1} = 0 \mid A_t = 1)$ , which is a straightforward computation assuming that the distribution of  $\tau$  is known. Hence, the process  $\{A_t\}$  is a nonhomogenous Markov chain. To reduce notation, we write  $\hat{Z}_t = (P_t, A_t)$ , the spot price and customer arrival indicator, as the *exogenous* part of the state variable. The full *state variable* for this problem is  $S_t = (R_t, \hat{Z}_t) = (R_t, P_t, A_t)$ , which takes values in the state space  $\mathcal{S} = \mathcal{R} \times \mathbb{R}_+ \times \{0, \pm 1\}$ . Throughout this paper, we use the notation  $\hat{Z}_{t+1}(s)$  to represent  $\hat{Z}_{t+1} \mid S_t = s$  and  $P_{t+1}(p)$  to represent  $P_{t+1} \mid P_t = p$ , in order to clarify certain expressions. We are now ready to define the cost function, which is given by

$$c_t(S_t, x_t, D_{t+1}) = (x_t P_t - f_t) \mathbf{1}_{\{t < \tau\}} + D_{t+1}(R_t, P_t) \mathbf{1}_{\{t = \tau\}} \in \mathcal{Z}_{t+1}. \quad (9)$$

The first term is the cost of energy before the arrival of the customer and the second term is the future cost of the customer at the time of arrival.

Finally, let  $\{X_0^\pi, X_1^\pi, \dots, X_T^\pi\}$  be a policy (indexed by  $\pi$ ) where  $X_t^\pi : \mathcal{S} \rightarrow [0, x_{\max}]$  is the decision function at time  $t$ . Following  $\pi$ , we produce a sequence of states  $S_0^\pi, S_1^\pi, \dots, S_T^\pi$  and a sequence of terminal value random variables  $D_1^\pi, D_2^\pi, \dots, D_{T+1}^\pi$ , thereby obtaining the costs  $C_t^\pi = c_{t-1}(S_{t-1}^\pi, X_{t-1}^\pi(S_{t-1}^\pi), D_t^\pi) \in \mathcal{Z}_t$ , for  $t = 1, 2, \dots, T$ . The risk-averse sequential decision problem that we aim to solve is given in (4).

Define the static risk measures  $\text{CVaR}_{\alpha_t} : \mathcal{Z} \rightarrow \mathbb{R}$  and  $\rho_{\beta_t} : \mathcal{Z} \rightarrow \mathbb{R}$ , where  $\mathcal{Z}$  is a space



of random variables,

$$\text{CVaR}_{\alpha_t}(X) = \inf_u \left\{ u + \frac{1}{1 - \alpha_t} \mathbf{E}[(X - u)^+] \right\}, \quad \rho_{\beta_t}(X) = (1 - \lambda_t) \mathbf{E}[X] + \lambda_t \text{CVaR}_{\alpha_t}(X).$$

We remark that when we write  $\text{CVaR}_{\alpha_t}$  without the subscript  $t$  as in (6), we are referring to the static case (the definition is analogous to (6), except with  $\mathbf{E}_t[\cdot]$  replaced with  $\mathbf{E}[\cdot]$ ). The result is a real number rather than a random variable, and the same holds for  $\rho_{t,\beta}$  versus  $\rho_{\beta_t}$ . Note that  $\rho_{\beta_t}$  is a *coherent risk measure*, as axiomatized in Artzner et al. [1999], meaning that it satisfies the properties of convexity, monotonicity, translation invariance, and positive homogeneity. The *value at risk* for some risk level  $\gamma$  of a random variable  $X$  is defined to be the  $\gamma$ -quantile, or  $\text{VaR}_\gamma(X) = \inf\{u : \mathbf{P}(X \leq u) > \gamma\}$ , and it is well-known that an equivalent representation of CVaR is (see, e.g., Pflug [2000], Acerbi and Tasche [2002])

$$\text{CVaR}_\gamma(X) = \frac{1}{1 - \gamma} \int_\gamma^1 \text{VaR}_w(X) dw. \quad (10)$$

Let  $\pi_\beta^*$  be the (index of) an optimal policy for (4) where  $\rho_t$  are specified using  $\beta_t$ . For convenience, we denote  $X_{t,\beta}^* = X_t^{\pi_\beta^*}$  to be the decision function at time  $t$ . The theorem below serves as the foundation for risk-averse MDPs; it is an analog of the well-known risk-neutral Bellman recursion, with expectations replaced by one-step risk measures (in our case, they are taken to be  $\rho_{\beta_t}$  at time  $t$ ). Since  $\beta$  includes  $\beta_t$  for all  $t$ , the subscript notation  $(t, \beta)$  allows the time-dependent risk parameter  $\beta_t$  to be used in the equations below; strictly speaking, it suffices to write  $(t, \beta_t)$ , but the notation becomes unwieldy.

**Theorem 1** (Risk-Averse Bellman Recursion, Ruszczyński [2010]).

Let  $\beta_t = (\lambda_t, \alpha_t)$  and suppose for any resource state  $r$  and spot price  $p$ , the value function satisfies  $V_{T,\beta}^*(r, p, 1) = V_{T,\beta}^*(r, p, 0) = \rho_{\beta_t}(D_{t+1}(r, p))$  and  $V_{T,\beta}^*(r, p, -1) = 0$  as a boundary condition. For each of the earlier time periods  $t < T$ ,

$$V_{t,\beta}^*(s) = \min_{x \in \mathcal{X}(r)} \rho_{\beta_t}(c_t(s, x, D_{t+1}) + V_{t+1,\beta}^*(S_{t+1}) \mid S_t = s),$$

for all  $s \in \mathcal{S}$ . The optimal risk-averse policy is given by

$$X_{t,\beta}^*(s) \in \arg \min_{x \in \mathcal{X}_t(r)} \rho_{\beta_t}(c_t(s, x, D_{t+1}) + V_{t+1,\beta}^*(S_{t+1}) \mid S_t = s),$$

i.e., it is greedy with respect to the optimal value function.

We now introduce some useful notation for states of the form  $s = (r, p, 1)$ , i.e., those states where the customer has not yet arrived. These are the “interesting” states, because the future evolution of the MDP is no longer affected by decisions when  $A_t \neq 1$ . First, it is useful for us to refer to a measure of risk on the future cost distribution. Specifically, we are interested in the value at risk and conditional value at risk of the future cost for states.



Hence, for  $s = (r, p, 1)$ , define the functions

$$v_{t,\alpha}^*(r, p) = \text{VaR}_{\alpha_t} [V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s))] \quad \text{and} \quad c_{t,\alpha}^*(r, p) = \text{CVaR}_{\alpha_t} [V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s))].$$

Also in the case of  $a = 1$ , it is helpful for us to define the *post-decision value function*  $\tilde{V}_{t,\beta}(r, p) = \rho_{\beta_t} [V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(r, p, 1))]$ , where for  $s = (r, p, 1)$ , we have the relationship

$$V_{t,\beta}^*(s) = \min_{x \in \mathcal{X}(r)} xp + \tilde{V}_{t,\beta}(r + x, p), \quad (11)$$

due to the translation invariance property of the coherent risk measure  $\rho_{\beta_t}$ .

### 2.3 Simple Properties of the Value Function

For the sake of completeness, we now state some useful and obvious facts regarding the optimal value function. The cost function is zero for any  $t > \tau$ , i.e., after the arrival of the customer, and therefore  $V_{t,\beta}^*(r, p, -1) = 0$  for all  $t, r$ , and  $p$ . We now turn our attention to  $t = \tau$ , in particular, the quantity  $V_{t,\beta}^*(r, p, 0) = \rho_{\beta_t}(D_{t+1}(r, p))$ . We can directly compute its value using the closed form representation of CVaR of a normal random variable given in [Rockafellar and Uryasev, 2000, Proposition 1], obtaining

$$\rho_{\beta_t}(D_{t+1}(r, p)) = m_t(r, p) + \lambda_t k(\alpha_t) \sigma_t(r, p) =: d_{t,\beta}(r, p),$$

where  $k(\alpha_t) = (1 - \alpha_t)^{-1} \phi(\Phi^{-1}(1 - \alpha_t))$ , with  $\phi$  and  $\Phi$  being the density and cumulative distribution function of a standard normal random variable, respectively. Hence,  $d_{t,\beta}(r, p)$  can be thought of as a “terminal cost” that can occur at any time  $t$  due to the fact that  $\tau$  is random. The following assumption provides some conditions on  $m_t$  and  $\sigma_t$ , the mean and standard deviation of  $D_{t+1}(r, p)$ . In this paper, we use the shorthand  $\partial_a = \frac{\partial}{\partial a}$  as the partial derivative operator for a variable  $a$ .

**Assumption 1.** Suppose the following statements regarding  $m_t(r, p)$  and  $\sigma_t(r, p)$  hold for each  $t$ .

- (i) The function  $d_{t,\beta}(r, p)$  is decreasing in  $r$ , increasing in  $p$ , and continuously differentiable in both  $r$  and  $p$ . Moreover,  $\partial_p d_{t,\beta}(r, p) \neq 0$  for all resource states  $r$  and spot prices  $p$ .
- (ii) The function  $d_{t,\beta}(r, p)$  is strictly convex in the resource  $r$  and the partial derivative  $\partial_r d_{t,\beta}(r, p)$  is independent of the spot price  $p$ .
- (iii) The function  $d_{t,\beta}(r, p)$  is Lipschitz with constant  $L_d$  in both  $r$  and  $p$ , i.e.,  $|d_{t,\beta}(r, p) - d_{t,\beta}(r', p')| \leq L_d \|(r, p) - (r', p')\|_2$ .

The rationale for Assumption 1(i) is that it is better to have more in storage (thus, the *cost* should decrease in both mean and variance). In addition, the future cost increases in  $p$  because higher prices today may signal higher prices in the future. The rationale for

Assumption 1(ii) is that the value of an additional unit of energy decreases in the view of the customer (as this is a measure of the customer's satisfaction) — this is a *diminishing marginal utility* assumption. The justification for the second half of (ii) is that the spot price should have no bearing on the customer's "marginal satisfaction" given an additional unit of storage in the vehicle. This is because, practically speaking, customers are unaware of the prices on the spot market. Assumption 1(iii) is a technical condition. Finally, note that since  $\mathbf{P}(A_T = 1) = 0$ , we never require the value  $V_{T,\beta}^*(r, p, 1)$  when computing the optimal policy. We somewhat arbitrarily set

$$V_{T,\beta}^*(r, p, 1) = V_{T,\beta}^*(r, p, 0) = d_{T,\beta}(r, p) \quad (12)$$

for consistency and simplicity of exposition.

### 3 The Risk-Averse Optimal Policy

In this section, we some structural properties of the risk-averse optimal policy and the associated optimal value function. We then introduce the notion that a spectrum of risk-averse policies can be generated by varying the risk parameter  $\beta$  and explain the usefulness of this idea in a practical setting. First, we show that the optimal value function is convex in the resource dimension  $r$  of the state variable, a useful property for determining the structure of the optimal policy.

**Proposition 1** (Convexity). *Let  $s = (r, p, 1) \in \mathcal{S}$  be the state variable. The following statements hold for every  $t$  and  $\beta_t$ : (i)  $xp + \tilde{V}_{t,\beta}(r + x, p)$  is strictly convex in  $x$  on  $\mathcal{X}(r)$ , (ii)  $\tilde{V}_{t,\beta}(s)$  is strictly convex in the resource  $r$ , and (iii)  $V_{t,\beta}^*(s)$  is strictly convex in the resource  $r$ .*

*Proof.* See Appendix A. □

The optimal value function also satisfies a monotonicity property, as shown in the following lemma. Note that monotonicity also holds in the resource state variable  $r$ , but it is not necessary for our subsequent analysis.

**Proposition 2** (Monotonicity in Spot Price). *Suppose the conditions of Assumption 1 are satisfied. For  $t \leq T$ ,  $r \in \mathcal{R}$ , and  $a \in \{0, 1\}$ , the optimal value function  $V_{t,\beta}^*(r, p, a)$  is (strictly) increasing in the spot price  $p$ .*

*Proof.* See Appendix A. □

The next lemma shows that part (iv) of Assumption 1 extends backwards for all  $t$ .

**Lemma 1** (Lipschitz Property). *For any  $r, r' \in \mathcal{R}$ ,  $p, p' \in \mathbb{R}$ , and  $a \in \{0, \pm 1\}$  and for each  $t \leq T$ , there exists deterministic  $L_t > 0$  such that the optimal value function  $V_{t,\beta}^*$  satisfies*

$$|V_{t,\beta}^*(r, p, a) - V_{t,\beta}^*(r', p', a)| \leq L_t \|(r, p) - (r', p')\|_2,$$

where  $\|\cdot\|_2$  is the Euclidean norm.

*Proof.* See Appendix A. □

In particular, Lemma 1 implies that the optimal value function is jointly continuous in the resource state  $r$  and the spot price  $p$ . It also allows us to interchange differentiation with expectation, as we will see in Lemma 4.

We now prove that the optimal policy is of the so called *order-up-to* type. In this result, the thresholds are given as a function of the risk parameters  $\beta$  and the spot price. The precise behavior of the thresholds are characterized in Section 4. Note that when  $a \neq 1$ , decisions no longer influence the evolution of the MDP and therefore we can arbitrarily set  $X_{t,\beta}^*(r, p, a) = 0$ . Hence, the optimal policy is interesting only when  $a = 1$  and to simplify notation, we write  $x_{t,\beta}^*(r, p) = X_{t,\beta}^*(r, p, 1)$ .

**Theorem 2** (Optimal Policy). *For each time  $t < T$  and spot price  $p$ , there exists a threshold resource level  $r_{t,\beta}^*(p)$  such that the optimal policy takes the following form:*

$$x_{t,\beta}^*(r, p) = \min\{r_{t,\beta}^*(p) - r, x_{\max}\} \mathbf{1}_{\{r \leq r_{t,\beta}^*(p)\}},$$

where  $r_{t,\beta}^*(p)$  is a value of  $\tilde{r}$  that minimizes the expression  $\tilde{r}p + \tilde{V}_{t,\beta}(\tilde{r}, p)$ . Furthermore, the threshold  $r_{t,\beta}^*(p)$  is continuous in  $p$  and the optimal policy  $x_{t,\beta}^*(r, p)$  is continuous in the resource and spot price pair  $(r, p)$ .

*Proof.* To obtain an equivalent formulation of (11) in terms of the post-decision resource state, we make a substitution  $\tilde{r} = r + x$  and  $x = \tilde{r} - r$ . Define  $\mathcal{R}(r) = \{\tilde{r} : r \leq \tilde{r} \leq \min\{r + x_{\max}, R_{\max}\}\}$  as the set of possible post-decision resource levels. We have

$$x_{t,\beta}^*(r, p) \in \left( \arg \min_{\tilde{r} \in \mathcal{R}(r)} \tilde{r}p + \tilde{V}_{t,\beta}(\tilde{r}, p) \right) - r. \quad (13)$$

Note that the objective is convex in  $\tilde{r}$  by Proposition 1. Since  $\mathcal{R}(r) \subseteq \mathcal{R}$ , consider the solution  $r_{t,\beta}^*(p)$  to the relaxed optimization problem

$$r_{t,\beta}^*(p) \in \arg \min_{\tilde{r} \in \mathcal{R}} \tilde{r}p + \tilde{V}_{t,\beta}(\tilde{r}, p). \quad (14)$$

Clearly, if  $r_{t,\beta}^*(p) \in \mathcal{R}(r)$ , then it is an optimal solution to the optimization problem within (13) and the optimal decision is  $r_{t,\beta}^*(p) - r$ . There are two ways in which  $r_{t,\beta}^*(p) \notin \mathcal{R}(r)$  can happen. First, if  $r_{t,\beta}^*(p) < r$ , then by convexity,  $\tilde{r}p + \tilde{V}_{t,\beta}(\tilde{r}, p)$  is nondecreasing on  $\mathcal{R}(r)$  and an optimal solution is  $\tilde{r} = r$ , implying an optimal decision of 0. The second case is when  $r + x_{\max} < r_{t,\beta}^*(p) \leq R_{\max}$ . Again, by convexity, it must be true that  $\tilde{r}p + \tilde{V}_{t,\beta}(\tilde{r}, p)$  is nonincreasing on  $\mathcal{R}(r)$  and an optimal solution is  $\tilde{r} = r + x_{\max}$ . This necessarily gives us an optimal decision of  $x_{\max}$ . Finally, combining these three cases allows us to conclude the structure of the optimal policy stated in the theorem. Now, notice that it follows by Lemma 1 and properties of  $\rho_{\beta_t}$  that the objective of (14) is continuous in  $(\tilde{r}, p)$ . Since the constraint set is compact, Berge's maximum theorem Berge [1963] applies and we can conclude that

the set of minimizers (i.e., the argmin) is upper hemicontinuous in  $p$  (see, e.g., [Stockey and Lucas, Jr. \[1989\]](#) for a definition). However, Proposition 1 implies that the minimum of (14) is attained at a single point and therefore, by [\[Border, 1989, Proposition 11.9\]](#),  $r_{t,\beta}^*(p)$  is continuous in  $p$ . The second conclusion of the theorem follows as  $x_{t,\beta}^*(r, p)$  is a continuous function of  $r_{t,\beta}^*(p) - r$ .  $\square$

By virtue of this theorem, the optimal policy can be described simply by its thresholds  $r_{t,\beta}^*(p)$  for any fixed  $\beta$ . The following definition formalizes the idea that we may vary the risk parameters  $\beta$  over some set  $\mathcal{B}$  to generate many risk-averse policies optimal with respect to various objectives.

**Definition 1.** The *spectrum of risk-averse policies over a set  $\mathcal{B}$*  is the set of all optimal risk-averse policies generated by risk parameters  $\beta \in \mathcal{B}$ . We denote it by  $\mathcal{R}^*(\mathcal{B}) = \{r_{t,\beta}^*(p) : t < T, p \in \mathbb{R}, \beta \in \mathcal{B}\}$ .

It is often the case in the literature that the issue of specifying the parameters of a risk measure (in our case,  $\beta$ ) is neglected or handled in an ad-hoc manner; however, risk preferences are not fully fixed until both the risk measure *and* its parameters are chosen. Therefore, an understanding of  $\mathcal{R}^*(\mathcal{B})$  can have important practical implications: if  $\mathcal{R}^*(\mathcal{B})$  were known, then a manager has the capability of choosing a policy (through simulations or other methods of selection) that matches her risk appetite, *with the assurance* that the policy is optimal under some instance (i.e., a choice of  $\beta$ ) of the risk-averse MDP given in (4), which can be thought of as a decision theoretic basis for the operating policy.

It is now natural to question the converse: would such an assurance be available if the manager were to heuristically choose threshold levels to control risk? In other words, given a policy by specifying thresholds for each  $p$  and  $t$ , is this policy necessarily the solution to (4) for *some* choice of  $\beta$ ? Intuitively, we expect to answer in the negative: it should not be the case that any sequence of thresholds is the solution to a complicated MDP. We cannot simply assume that a heuristically tuned sequence of thresholds is optimal under some risk-averse MDP specified by  $\beta$ .

Indeed, it is easy to construct examples where  $r_{t,\beta}^*(p)$  is independent of  $\beta$ . For example, suppose parameters of the future cost distribution can be written  $m_t(r, p) = m_t(r)$  and  $\sigma_t(r, p) = \sigma_t(p)$ , so that  $d_{T,\beta}(r, p) = m_t(r) + \lambda_T k(\alpha_T) \sigma_t(p)$ . Using  $\mathbf{P}(A_T = 1) = 0$  and (14), we have  $r_{T-1,\beta}^*(p) \in \arg \min_{\tilde{r} \in \mathcal{R}} \tilde{r} p + m_t(\tilde{r}) + \rho_{\beta_{T-1}} [\lambda_T k(\alpha_T) \sigma_t(P_T(p))]$ , which clearly does not depend on  $\beta$ . Therefore, it is in fact useful for us to understand the properties of  $\mathcal{R}^*(\mathcal{B})$  in order to control risk in a theoretically justified manner. See Figure 1 for an illustration of the relationships described here. In the next section, we characterize the structure of  $r_{t,\beta}^*(p)$  in  $\beta$  and  $p$ , and in Section 5, we offer a computational method for approximating  $\mathcal{R}^*(\mathcal{B})$ .

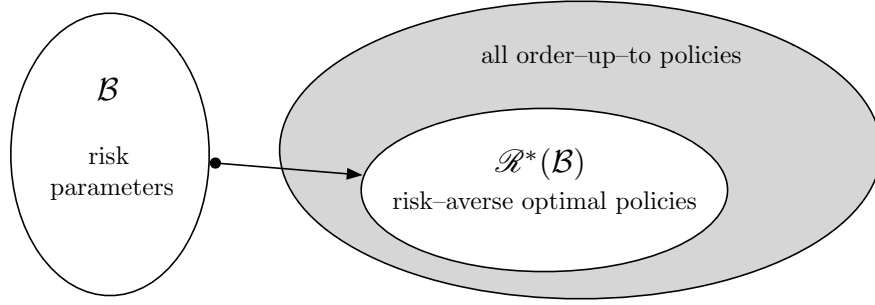


Figure 1: Spectrum of Risk-Averse Policies

## 4 Behavior of Optimal Threshold Levels

So far, we have seen the structure of the optimal policy for fixed risk parameters  $\beta$  and a fixed spot price  $p$ . It is useful, however, to understand how  $r_{t,\beta}^*(p)$  moves with  $\beta_t$  and  $p$ . The main result of this paper is given in Theorem 3, where we show that the map  $(\lambda_t, \alpha_t, p) \rightarrow r_{t,\beta}^*(p)$  is nondecreasing in  $\lambda_t$  and  $\alpha_t$  and nonincreasing in the spot price  $p$ .

In particular, this allows us to draw a connection between the optimal risk-neutral policy, i.e., when  $\lambda_t = 0$ , and an optimal risk-averse policy, i.e., for some  $\lambda_t > 0$ . To the best of our knowledge, the question of characterizing the relationship between optimal risk-neutral policies and policies formed under the dynamic risk measure framework has not been studied in the literature; in this paper, we present some results for our specific model, but leave a more general study to future work.

### 4.1 Analysis of Optimal Thresholds

To properly analyze the behavior of the thresholds, we need several lemmas to cover the following: (1) differentiability of the value at risk of the optimal value function, (2) relating differentiability of the post-decision value function to differentiability of the pre-decision value function, and (3) differentiability of CVaR.

We begin with some general properties. Applying the monotonicity property in  $p$  given in Proposition 2, we know that  $V_{t+1,\beta}^*(r, p, 0)$  and  $V_{t+1,\beta}^*(r, p, 1)$  are invertible in  $p$ . Denote the inverse functions by  $\nu_{t+1}^0(r, p)$  and  $\nu_{t+1}^1(r, p)$ , respectively, so that for each  $r$  and  $p$ ,

$$\nu_{t+1}^0(r, V_{t+1,\beta}^*(r, p, 0)) = p \quad \text{and} \quad \nu_{t+1}^1(r, V_{t+1,\beta}^*(r, p, 1)) = p. \quad (15)$$

For  $a \in \{0, 1\}$ , the inverse function  $\nu_{t+1}^a(r, p)$  is differentiable in  $p$  when  $\partial_p V_{t+1,\beta}^*(r, p, a)$  exists and is nonzero; by an elementary calculus result, the “inverse function theorem,” we have

$$\partial_p \nu_{t+1}^a(r, p) = (\partial_p V_{t+1,\beta}^*(r, p, a))^{-1}. \quad (16)$$

On a similar note, using implicit differentiation on (15), we see that  $\nu_{t+1}^a(r, p)$  for  $a \in \{0, 1\}$  is differentiable in  $r$  when  $\partial_r V_{t+1, \beta}^*(r, p, a)$  exists,  $\partial_p V_{t+1, \beta}^*(r, p, a)$  exists, and  $\partial_p V_{t+1, \beta}^*(r, p, a)$  is nonzero; it is given by the formula

$$\partial_r \nu_{t+1}^a(r, p) = -\frac{\partial_r V_{t+1, \beta}^*(r, \nu_{t+1}^a(r, p), a)}{\partial_p V_{t+1, \beta}^*(r, \nu_{t+1}^a(r, p), a)}. \quad (17)$$

We remark that the term in the denominator is written with an abuse of notation: to be precise, we mean the derivative with respect to  $p$  evaluated at  $\nu_{t+1}^a$ , i.e.,  $\partial_p V_{t+1, \beta}^*(r, p, a)|_{p=\nu_{t+1}^a}$ . Throughout this paper, differentiation in a particular dimension of a function is written with respect to the variable used in the function's original definition.

We now derive the distribution function of the random variable  $V_{t+1, \beta}^*(r, \hat{Z}_{t+1}(s))$  when  $s = (r, p, 1)$ , i.e., the future value starting from state  $s$ . This becomes useful in the ensuing analysis. Recall that  $P_{t+1} | P_t = p$  has the same distribution as  $\mu_P(p) + \epsilon_{v_P}$ . Let  $\Phi_{v_P}$  be the distribution function of a mean zero normal random variable with variance  $v_P$ . Also, let  $q_t^0 = \mathbf{P}(A_{t+1} = 0 | A_t = 1)$  and  $q_t^1 = \mathbf{P}(A_{t+1} = 1 | A_t = 1)$  be the transition probabilities from the state  $A_t = 1$ . Then, it is easy to show that the distribution function of  $V_{t+1, \beta}^*(r, \hat{Z}_{t+1}(s))$  can be written as

$$F_{t+1}^V(r, p, \theta) = \mathbf{P}(V_{t+1, \beta}^*(r, \hat{Z}_{t+1}(s)) \leq \theta) = 1 - \sum_{a \in \{0, 1\}} q_t^a \Phi_{v_P}(\nu_{t+1}^a(r, \theta) - \mu(p)). \quad (18)$$

Let  $\phi_{v_P}$  be the density of a mean zero normal random variable with variance  $v_P$ . Differentiating and using the previous expressions for  $\partial_p \nu_{t+1}^a(r, p)$  and  $\partial_r \nu_{t+1}^a(r, p)$ , we obtain

$$\begin{aligned} \partial_r F_{t+1}^V(r, p, \theta) &= \sum_{a \in \{0, 1\}} \left[ q_t^a \phi_v(\nu_{t+1}^a(r, \theta) - \mu(p)) \frac{\partial_r V_{t+1, \beta}^*(r, \nu_{t+1}^a(r, \theta), a)}{\partial_p V_{t+1, \beta}^*(r, \nu_{t+1}^a(r, \theta), a)} \right], \\ \partial_p F_{t+1}^V(r, p, \theta) &= \sum_{a \in \{0, 1\}} \left[ q_t^a \phi_v(\nu_{t+1}^a(r, \theta) - \mu(p)) e^{-\kappa_P} \right], \\ \partial_\theta F_{t+1}^V(r, p, \theta) &= - \sum_{a \in \{0, 1\}} \left[ q_t^a \phi_v(\nu_{t+1}^a(r, \theta) - \mu(p)) (\partial_p V_{t+1, \beta}^*(r, \theta, a))^{-1} \right]. \end{aligned}$$

Hence, existence of these derivatives depends on the existence of the partial derivatives of  $V_{t+1, \beta}^*$ . Next, since  $\Phi_v(\cdot)$  and  $\nu_{t+1}^a(r, \cdot)$  are continuous and increasing functions, we know that  $F_{t+1}^V$  is continuous and increasing in  $\theta$ . Thus, the value at risk, or quantile function,  $v_{t, \alpha}^*(r, p)$  is simply the inverse of  $F_{t+1}^V(r, p, \cdot)$  evaluated at  $\alpha_t$ . In particular, we must have  $F_{t+1}^V(r, p, v_{t, \alpha}^*(r, p)) = \alpha_t$  for any resource state  $r$  and spot price  $p$ .

**Lemma 2** (Differentiability of Value at Risk). *Suppose  $V_{t+1, \beta}^*$  is differentiable in both  $r$  and  $p$ . Then, the derivatives of  $v_{t, \alpha}^*(r, p)$  are given by the expressions*

$$\partial_r v_{t, \alpha}^*(r, p) = -\frac{\partial_r F_{t+1}^V(r, p, v_{t, \alpha}^*(r, p))}{\partial_\theta F_{t+1}^V(r, p, v_{t, \alpha}^*(r, p))} \quad \text{and} \quad \partial_p v_{t, \alpha}^*(r, p) = -\frac{\partial_p F_{t+1}^V(r, p, v_{t, \alpha}^*(r, p))}{\partial_\theta F_{t+1}^V(r, p, v_{t, \alpha}^*(r, p))}.$$

*Proof.* These expressions follow by implicit differentiation of  $F_{t+1}^V(r, p, v_{t,\alpha}^*(r, p)) = \alpha_t$ .  $\square$

The next lemma relates differentiability of the post-decision value function to differentiability in the pre-decision value function. It also verifies that when the post-decision value function has a nonzero derivative, then the same holds for the pre-decision value function. This fact allows us to compute (16) and (17) whenever the derivatives exist.

**Lemma 3** (Differentiability of Value Function in Spot Price). *Suppose that for some  $t$  and resource  $r$ , the partial derivative  $\partial_p \tilde{V}_{t,\beta}(r, p)$  exists for any spot price  $p$  and is nonzero. Then, it follows that  $\partial_p V_{t,\beta}^*(r, p, 1)$  also exists for any spot price  $p$  and is nonzero.*

*Proof.* See Appendix A.  $\square$

The notion of *piecewise continuously differentiability* for some function  $f : [a, b] \rightarrow \mathbb{R}$  states that there exists  $a_i$  with  $a = a_0 < a_1 < \dots < a_n = b$  such that  $f$  is continuously differentiable when restricted to  $[a_i, a_{i+1}]$  for each  $i < n$ . More precisely, for a fixed  $i$ ,  $f'(x)$  at  $x \in (a_i, a_{i+1})$  is the derivative, while  $f'(a_i)$  is the right derivative, and  $f'(a_{i+1})$  is the left derivative and  $f'$  is continuous on  $[a_i, a_{i+1}]$ . By Theorem 2, the optimal policy  $x_{t,\beta}^*(r, p)$  is piecewise linear and hence, piecewise continuously differentiable in  $r$ .

**Assumption 2.** Suppose that: if for some  $t$ , the post-decision value function  $\tilde{V}_{t,\beta}(r, p)$  is piecewise continuously differentiable in  $r$ , then  $V_{t,\beta}^*(r, p, 1)$  is also piecewise continuously differentiable in  $r$ .

Assumption 2 implies that we do not introduce intervals of non-differentiability. To understand it further (and why it is not a restrictive assumption), note that by Theorem 2, we can represent the optimal value function at  $s = (r, p, 1)$  by substituting in the optimal policy, so that we have

$$V_{t,\beta}^*(s) = x_{t,\beta}^*(r, p)p + \tilde{V}_{t,\beta}(r + x_{t,\beta}^*(r, p), p). \quad (19)$$

The function  $r \mapsto r + x_{t,\beta}^*(r, p)$  takes a constant value (denote it  $R$ ) on the interval  $(r_{t,\beta}^*(p) - x_{\max}, r_{t,\beta}^*(p))$ . Note that for fixed  $p$ , the function  $xp$  is linear in  $x$  and the policy  $x_{t,\beta}^*(r, p)$  is piecewise linear in  $r$  with two breakpoints at  $r = x_{\max}$  and  $r = r_{t,\beta}^*(p)$ . Therefore, if we differentiate in  $r$ , we can see that for Assumption 2 (i) to hold, the value  $R$  must not coincide with a non-differentiability point of  $\tilde{V}_{t,\beta}$ .

The remaining missing link is to translate differentiability in the optimal value function  $V_{t+1,\beta}^*$  to differentiability in the post-decision value function  $\tilde{V}_{t,\beta}$ . The next lemma allows interchanging differentiation with the risk measure  $\rho_{\beta_t}$ . Its proof crucially depends on sensitivity analysis results for conditional value at risk due to Hong and Liu [2009].

**Lemma 4** (Interchanging Derivative and Risk Measure). *We consider differentiation in  $r$  and  $p$  separately, starting with  $r$ . Fix a spot price  $p$  and let  $s = (r, p, 1)$ . Let  $\bar{\mathcal{R}} \subseteq \mathcal{R}$  be a*



set of resource values where  $V_{t+1,\beta}^*(r, p, 1)$  is differentiable. Then, for every  $r \in \bar{\mathcal{R}}$ ,

$$\begin{aligned} \partial_r \tilde{V}_{t,\beta}(r, p) &= (1 - \lambda_t) \mathbf{E}[\partial_r V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s))] \\ &\quad + \lambda_t \mathbf{E}[\partial_r V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s)) \mid V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s)) \geq v_{t,\alpha}^*(r, p)]. \end{aligned}$$

Now, we fix a resource state  $r$ . Suppose that  $V_{t+1,\beta}^*(r, p, 1)$  is differentiable everywhere in the spot price. It follows that for every  $p \in \mathbb{R}$ ,

$$\begin{aligned} \partial_p \tilde{V}_{t,\beta}(r, p) &= (1 - \lambda_t) \mathbf{E}[\partial_p V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s))] \\ &\quad + \lambda_t \mathbf{E}[\partial_p V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s)) \mid V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s)) \geq v_{t,\alpha}^*(r, p)]. \end{aligned}$$

We remark that “interchange” is used loosely here; the second term (in both cases above) resembles the “CVaR of the derivative” but note that the conditioning event has not changed.

*Proof.* See Appendix A. □

We are almost ready to state and prove the main result, which describes the behavior of the thresholds  $r_{t,\beta}^*(p)$  with respect to the risk parameters  $\beta_t$  and the spot price  $p$ . To analyze the charging threshold  $r_{t,\beta}^*(p)$ , the optimization problem in question is

$$\min_{\tilde{r} \in \mathcal{R}} \tilde{r} p + \tilde{V}_{t,\beta}(\tilde{r}, p). \quad (20)$$

Because the post-decision value function is not necessarily differentiable for all  $r$ , we introduce a separate notation for the left-derivative

$$\tilde{V}'_{t,\beta}(r, p) = \lim_{h \downarrow 0} \frac{\tilde{V}_{t,\beta}(r, p) - \tilde{V}_{t,\beta}(r - h, p)}{h},$$

for each  $r$  in the interior of  $\mathcal{R}$ . This one-sided derivative exists by the convexity given in Proposition 1. Of course, if  $\tilde{V}_{t,\beta}(r, p)$  is differentiable at  $r$ , then  $\tilde{V}'_{t,\beta}(r, p)$  coincides with  $\partial_r \tilde{V}_{t,\beta}(r, p)$ . We extend the definition to the boundary by setting  $\tilde{V}'_{t,\beta}(0, p) = \lim_{r \downarrow 0} \tilde{V}'_{t,\beta}(r, p)$  and  $\tilde{V}'_{t,\beta}(R_{\max}, p) = \lim_{r \uparrow R_{\max}} \tilde{V}'_{t,\beta}(r, p)$ .

**Theorem 3** (Threshold Behavior). *Under Assumptions 1–2, the following statements regarding the behavior of the optimal threshold  $r_{t,\beta}^*(p)$  hold.*

- (i) *The threshold  $r_{t,\beta}^*(p)$  is nondecreasing in the risk parameters  $\beta_t = (\lambda_t, \alpha_t)$ . In other words, as risk aversion increases, the thresholds relax.*
- (ii) *The threshold  $r_{t,\beta}^*(p)$  is nonincreasing in the spot price  $p$ .*

*Proof.* See Appendix A. □

The proof is long, so we provide only a sketch here in the body of the paper. Before continuing, however, we make a few remarks. Considering that we are focused on risk-averse policies in this work, part (i) of Theorem 3 is the main and most interesting result.

See Figure 2 for a graphical depiction of the change in the policy as risk aversion increases; note that by Theorem 2, the shape of the policy does not change. To our knowledge, the relationship between risk-averse and risk-neutral policies has not been studied in the literature, and we believe that this analysis opens the door to a more general study of the effects of risk-aversion on the structure of optimal policies. Even though part (ii) of Theorem 3 seems intuitive, it depends heavily on the properties of the price process, which we exploit in the proof; see Secomandi [2010], where a similar result for a commodity trading problem is derived.

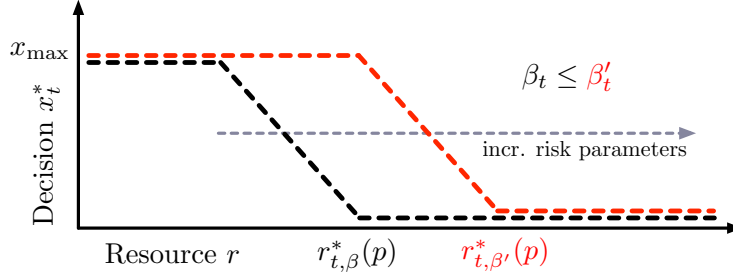


Figure 2: Illustration of Theorem 3 (i)

*Sketch of Proof of Theorem 3.* The proof is by backward induction on  $t$ , with a multi-part induction hypothesis, enumerated below:

- (a) For every spot price  $p$ , the value function  $\tilde{V}_{t,\beta}$  is piecewise continuously differentiable.
- (b) For every resource state  $r$ , the optimal value function  $\tilde{V}_{t,\beta}$  is differentiable in  $p$  at every spot price  $p$  and the derivative is nonzero.
- (c) The quantity  $p + \mathbf{E}[\partial_r V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s)) | V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s)) \geq v_{t,\alpha}^*(r, p)]$  is nondecreasing in the spot price  $p$  whenever the derivative in  $r$  exists.
- (d) The quantity  $p + \tilde{V}'_{t,\beta}(r, p)$  is nondecreasing in the spot price  $p$ . This implies part (ii) of the theorem.
- (e) The derivative of the value function  $\tilde{V}'_{t,\beta}(r, p)$  is nonincreasing in the spot price  $p$ .
- (f) The derivative of the value function  $\tilde{V}'_{t,\beta}(r, p)$  is componentwise nonincreasing in  $\beta_t = (\lambda_t, \alpha_t)$ . This implies part (i) of the theorem.

The general proof technique is to analyze derivatives and applying Lemma 4. Properties of  $\tilde{V}'_{t,\beta}$  are used to infer properties of the derivative of  $V_{t,\beta}^*$ , and finally, Lemma 4 is applied in order to obtain properties of  $\tilde{V}'_{t-1,\beta}$ . Part (c) at time  $t+1$  is used to prove parts (c) and (d) at time  $t$ , from which we conclude part (ii). This part of the proof uses a similar line of reasoning given in Secomandi [2010]. Part (e) at time  $t+1$  is used to prove parts (e) and (f) at time  $t$ , which gives us part (i) of the main result. Finally, parts (a) and (b) are technicalities needed for differentiation and are easy to argue.  $\square$

## 4.2 Approximating the Spectrum of Risk-Averse Policies

Consider the set  $\mathcal{B} = \{\beta : \beta_0 = \beta_1 = \dots = \beta_{T-1}\}$ , the set of risk parameters that are time-independent. A small abuse of notation allows us to consider  $\rho_{0,T-1}$  specified by composing  $\rho_\beta$ , with  $\beta$  varying over the set  $\mathcal{B} = \{\beta = (\lambda, \alpha) : \lambda \in [0, 1], \alpha \in (0, 1)\}$ . We now propose a methodology using polynomial optimization and sum of squares constraints to approximate  $\mathcal{R}^*(\mathcal{B})$ , given that the optimal policy has been computed for some values of  $\lambda, \alpha, p$ ; assume that for each  $t$ , we observe a threshold of  $r_t^i$  at  $(\beta^i, p^i)$  for  $i = 1, 2, \dots, N$ . Let  $\bar{r}_{t,\beta}(p)$  be the approximation to  $r_{t,\beta}^*(p)$  and recall that  $\beta = (\lambda, \alpha)$ . The approximate spectrum of policies is analogously denoted  $\bar{\mathcal{R}}(\mathcal{B}) = \{\bar{r}_{t,\beta}(p) : t < T, p \in \mathbb{R}, \beta \in \mathcal{B}\}$ .

The main idea is to take advantage of Theorem 3 and employ a form of shape constrained regression, as shown in the following  $l_1$  norm regression subject to monotonicity constraints:

$$\begin{aligned} & \text{minimize} && \sum_i |\bar{r}_{t,\beta^i}(p^i) - r_t^i| \\ & \text{subject to} && \bar{r}_{t,\beta}(p) \text{ nondecreasing in } \lambda, \alpha, \\ & && \bar{r}_{t,\beta}(p) \text{ nonincreasing in } p. \end{aligned}$$

Many monotone regression techniques are specified only for the one dimensional case [Mukerjee, 1988, Dette et al., 2006], but following the general idea given in Ahmadi and Majumdar [2014] for convex regression, if we restrict  $\bar{r}_t$  to take the form of a polynomial, then the monotonicity constraints can be approximated with *sum of squares* (SOS) constraints. The resulting problem can be solved efficiently using semi-definite programming [Parrilo, 2003]:

$$\begin{aligned} & \text{minimize}_{\text{poly } \bar{r}_t} && \sum_i |\bar{r}_{t,\beta^i}(p^i) - r_t^i| \\ & \text{subject to} && \partial_\lambda \bar{r}_{t,\beta}(p), \partial_\alpha \bar{r}_{t,\beta}(p) \text{ are SOS,} \\ & && -\partial_p \bar{r}_{t,\beta}(p) \text{ is SOS.} \end{aligned}$$

An additional advantage that a polynomial approximation of  $\bar{r}_{t,\beta}$  allows is that policy evaluation becomes computationally very cheap, allowing for extensive simulation and policy testing. We illustrate these ideas and their effectiveness in Section 5.

## 4.3 Practical Metrics for Risk and Reward

Despite the appealing theoretical properties, it is likely that the value function associated with dynamic risk measure such as  $\rho_{0,T-1}$ , i.e., the quantity  $V_{0,\beta}^*(S_0)$ , would be difficult to adopt as a *metric* for risk/reward in practice, due to the lack of a simple interpretation. This is unlike a risk-neutral framework, where the value function at time zero is simply interpreted as the expected cost of the policy. A manager of EV charging stations may *implement* policies generated by our risk-averse framework, while electing to *measure the performance* of these policies using metrics that have a direct relationship to the cash flow of the firm. In fact, it is quite common even in the academic literature to evaluate risk-averse

policies using a different objective function; oftentimes, the policies are evaluated along two separate metrics, risk and reward. For example, see [Philpott and de Matos \[2012\]](#), [Shapiro et al. \[2013\]](#), and [Çavus and Ruszczyński \[2014\]](#), where a risk-neutral objective is used for evaluation of reward while another metric is used for risk. We now formalize the connection between policies optimized via the dynamic risk measure framework and several practical metrics for risk and reward by stating a simple corollary of Theorem 3. The most natural metric for reward is, of course, to take the negative expectation of the spot market payments:

$$\text{Reward}(\pi) = \mathbf{E} \left[ - \sum_{t=0}^{\tau-1} X_t^\pi(S_t^\pi) P_t \right].$$

Two practical metrics of risk are (1)  $\text{Risk}_F^\delta(\pi)$ , the probability that a customer leaves without a nearly full charge and (2)  $\text{Risk}_D(\pi)$ , the expected future cost (lost future sales) of the customer, which can be respectively expressed as

$$\text{Risk}^F(\pi; \delta) = \mathbf{P}\{R_\tau^\pi < (1 - \delta) R_{\max}\} \quad \text{and} \quad \text{Risk}^D(\pi) = \mathbf{E}[D_{\tau+1}(R_\tau^\pi, P_\tau)],$$

where  $R_\tau^\pi$  is the amount of charge received by the customer at time  $\tau$  following policy  $\pi$ .  $\text{Risk}^F(\pi; \delta)$  represents “service reliability at some threshold,” while in the case of  $\text{Risk}^D(\pi)$ , we are also concerned with the amount of undercharge/lost future value, while this information would not be seen in  $\text{Risk}^F(\pi; \delta)$ . These examples have the characteristic that the risk of a policy is separate from the reward of a policy, mirroring the typical approach in industrial settings. In some sense, this is contrary to using the value function  $V_{0,\beta}^*(S_0)$ , where the risk and reward are both “encoded” into  $\rho_{0,T-1}$  through  $\lambda_t$  and  $\alpha_t$ .

**Corollary 4** (Behavior of Practical Metrics). *Recall that  $\pi_\beta^*$  is an optimal policy for (4) when the parameters of the dynamic risk measure  $\rho_{0,T-1}$  are specified using  $\beta$ . Under Assumptions 1–2, it follows that  $\beta \mapsto \text{Risk}^F(\pi_\beta^*; \delta)$ ,  $\beta \mapsto \text{Risk}^D(\pi_\beta^*)$ , and  $\beta \mapsto \text{Reward}(\pi_\beta^*)$  are nonincreasing.*

This result follows directly from the structure of the optimal policy given in Theorem 3. It is interesting because although we solve the risk-averse decision making problem given in (4) under a time-consistent dynamic risk measure, it is simultaneously true that practical metrics (that may otherwise be used in designing ad-hoc, heuristic solutions) vary in a way that is intuitively appealing. In other words, a manager may take comfort in the fact that the theoretically sound formulation of (4) may be used to generate a family of policies with varying risk versus reward trade-offs, both measured using practical metrics. We shed further light onto this issue in the numerical work in Section 5.

## 5 Numerical Results and Case Study

In this section, we investigate our results on a realistic instance of the EV charging problem. Using the polynomial optimization technique for approximating the spectrum of risk-averse policies, we illustrate the policies on the Pareto frontier when considering the competing objectives of maximizing  $\text{Reward}(\pi_\beta^*)$  while minimizing  $\text{Risk}^F(\pi_\beta^*; \delta)$  or  $\text{Risk}^D(\pi_\beta^*)$ .

### 5.1 Model Parameters

The spot market in our model is based on the CAISO real-time market, specifically the fifteen-minute market. Hence,  $t$  is measured in units of fifteen minutes and we choose a horizon of  $T = 20$  (or 4 hours). We calibrate an Ornstein–Uhlenbeck model using maximum likelihood estimation on CAISO data (the BAYSHOR2\_1\_N001 node) from February 1, 2016 to April 20, 2016 (over 7000 data points), resulting in the values  $\theta_P = 21.97$ ,  $\kappa_P = 0.55$ , and  $\sigma_P = 17.02$  for the spot price process  $P_t$ , measured in \$/MWh. The initial price is chosen to be  $P_0 = 21$  and the battery is assumed to be initially empty:  $R_0 = 0$ . The capacity of the battery is  $R_{\max} = 60$  kWh, in line with the entry version of the Tesla Model S from 2012 to 2015. This is a reasonable choice for a medium sized battery, as there are both EVs with smaller capacities (e.g., Nissan Leaf or BMW i3) and EVs with larger capacities (e.g., newer and higher end Tesla models).

We assume a 50 kW DC fast charging station, approximately corresponding to  $x_{\max} = 12$  kWh in a 15 minute interval. The time-based fee is set at \$2.00 per hour (an estimate obtained from ChargePoint), corresponding to  $f_t = 0.5$  for all  $t$ . The parameters of the terminal cost distribution are chosen to be  $m_t(r, p) = 9.88 e^{-5r/R_{\max}} + p/1000$  and  $\sigma_t(r, p) = R_{\max} - r + 1$ . We roughly estimate the “value of the customer” to be 9.88, or 10 visits (in, say, the next month) at  $(f_t \times 60 \text{ kWh}/x_{\max} - P_0/1000 \times 1.2 \times 60 \text{ kWh})$  \$/visit, where “1.2” is a risk-adjustment factor. It is often recommended to charge EV batteries up to 80%<sup>1</sup> — the exponent in  $m_t(r, p)$  is chosen so that  $m_t(r, p)$  decreases 98–99% from  $r = 0$  to  $r = 0.8 R_{\max}$ . Lastly, the distribution of  $\tau$  is constructed by discretizing two normal distributions (means of 6 and 12 with standard deviation 2) according to the method of Roy [2003] and creating a mixture between the two at probabilities 0.3 and 0.7. This can be thought of as mixing the behaviors of two types of customers, one that arrives promptly and one that arrives with some increased delay. The probabilities of arrival for  $t = 0, 1, \dots, 5$  are then set to zero, as the earliest customer return time is five charging intervals (60 kWh at 12 kWh/15 min). Plots of  $m_t(r, 0)$  and the distribution of  $\tau$  are given in Figure 3.

### 5.2 Exact MDP Results

In order to properly solve the risk-averse MDPs, discretization is necessary. We again use the method of Roy [2003] to discretize  $\epsilon_v$  to only contain integer outcomes. Ignoring the

<sup>1</sup><https://www.teslamotors.com/supercharger>

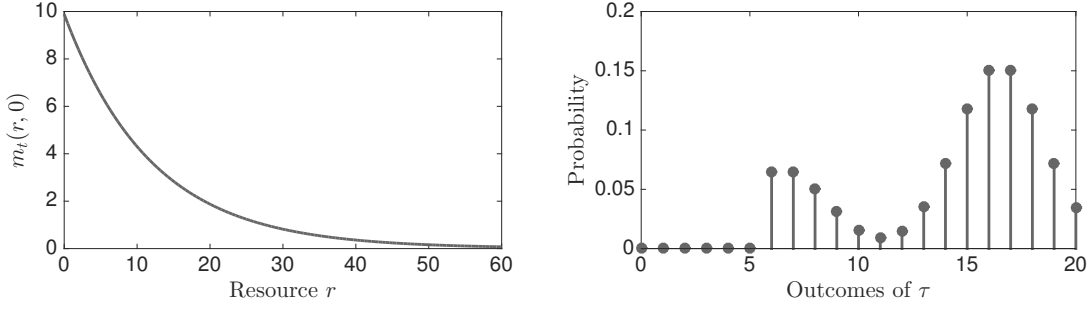


Figure 3: Model Parameters

outcomes with probability less than  $10^{-5}$ , we obtain a finitely supported version of the spot price process  $P_t$  taking values in the set  $\{-84, -83, \dots, 213\}$ . Such a discretization allows us to utilize the linear programming method for computing the CVaR, as pointed out in [Rockafellar and Uryasev \[2000\]](#). The resource state is finely discretized at a resolution of 0.5, so that  $R_t$  takes values in  $\{0, 0.5, \dots, R_{\max}\}$ . Each MDP contains a total of approximately 515,000 states. Expending a significant amount of CPU time, we solve a total of 399 MDPs for risk parameter values of  $\lambda \in \{0, 0.05, \dots, 1\}$  and values of  $\alpha \in \{0.05, 0.1, \dots, 0.95\}$ . The thresholds  $r_{t,\beta}^*(p)$  are then computed via (14).

We report our results using the illustration given in Figure 4 (the viewing angles of the two figures are different in order to show the structure of the data), serving also to experimentally verify the statements of Theorem 3. For example, we see in Figure 4A that the risk parameters  $(\lambda, \alpha)$  can dramatically change the policy: in the risk-neutral case,  $r_{t,\beta}^*(p) \approx 20$ , while in the most risk-averse case,  $r_{t,\beta}^*(p) \approx 60 = R_{\max}$  (meaning we should charge as much as possible). Figure 4B shows the second part of Theorem 3, that the thresholds are nonincreasing in  $p$ . In this case, we notice that at the same risk-level, a \$30 change in spot price can alter the threshold by as much as 35 kWh (when  $\lambda = 0$ ,  $\alpha = 0.05$ ). Notice that the viewing angles of the two figures are different in order to show the structure of the data.

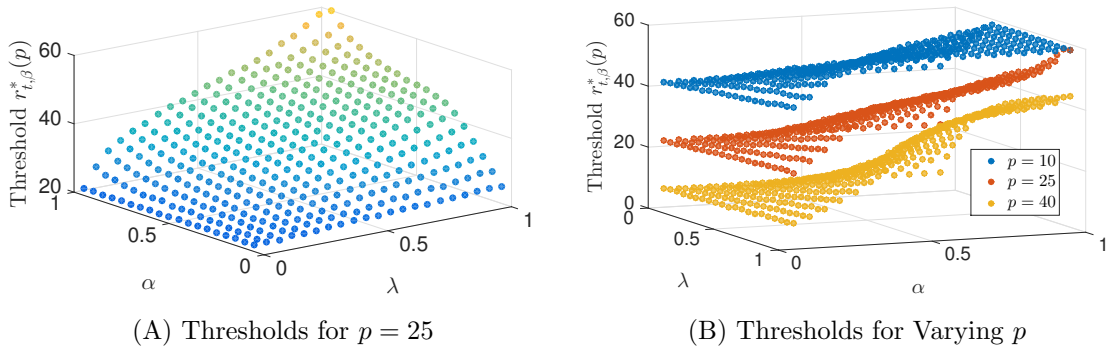


Figure 4: Behavior of Thresholds  $r_{t,\beta}^*(p)$  for  $t = 4$

### 5.3 Benchmarking the Approximate Spectrum of Risk-Averse Policies

As we previously mentioned, it is likely that the parameters  $(\lambda, \alpha)$  that produce satisfactory operating policies are not known in advance, partly due to the possibility that the firm operating the EV charging stations requires constraints on other metrics of risk and reward, such as those stated in Corollary 4, to be satisfied. Let us emphasize once again that it is not practical to compute the optimal policy for all possible parameters  $(\lambda, \alpha)$ : in our experiments, the compute time for a single MDP (or a single point in Figure 4A) is approximately 20 minutes on a 2.6 GHz machine using CPLEX (a state of the art solver) for linear programming. In order to produce the plots in Figure 4, a large and impractical amount of parallelization on a cluster is used. We refer to these as “benchmark policies,” on which we can test our approximations.

Thus, in this section we implement the procedure of using polynomial optimization with SOS constraints as outlined in Section 4.2. To benchmark the method, we fit the regression model on 36 optimal policies, fewer than 10% of the 399 MDPs used above: specifically, we assume that optimal policies for the parameters  $\lambda \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$  and  $\alpha \in \{0.05, 0.2, 0.4, 0.6, 0.8, 0.95\}$  are known and use them as observations for the regression model. We elect to use a degree 5 polynomial representation of the optimal policy, given by  $\bar{r}_{t,\beta}(p) = a_t^\top \phi(\lambda, \alpha, p)$ , where  $\phi(\lambda, \alpha, p)$  is the vector of all degree 5 monomials (56 terms) and  $a_t \in \mathbb{R}^{56}$  is the coefficient vector:

$$\phi(\lambda, \alpha, p) = [1, \lambda, \alpha, p, \lambda\alpha, \lambda p, \alpha p, \lambda^2, \alpha^2, p^2, \dots, \lambda^5, \alpha^5, p^5].$$

Using the SPOT toolbox [Megretski, 2013] and the MOSEK solver, the regression procedure to compute the coefficients  $a_t \in \mathbb{R}^{56}$  requires 7 seconds of CPU time for each  $t$ . Hence, the approximate spectrum of policies,  $\bar{\mathcal{R}}(\mathcal{B})$ , takes around 2.5 minutes of CPU time to produce.

In Figure 5, we compare the polynomial approximation (the surface) with the benchmarks (red dots) that we computed earlier. We remark that although we show all of the benchmark policies in Figure 5, only a small subset of 36 policies are actually used in the regression procedure. The SOS constraints indeed produce approximations that obey the structural results of Theorem 3.

In addition to the graphical comparison, a numerical one is necessary to show the effectiveness of the approximation procedure. Let  $\{X_0^\pi, X_1^\pi, \dots, X_{T-1}^\pi\}$  be a policy and let  $S_{t+1}^\pi = (R_t + X_t^\pi(S_t^\pi), P_{t+1}, A_{t+1})$ , with  $S_0^\pi = S_0$ . For a fixed  $\beta$ , the *value of the policy*, or the objective of the optimization (4), can be found by solving the recursive equations given by

$$\begin{aligned} V_t^\pi(s) &= \rho_\beta(c_t(s, X_t^\pi(s), D_{t+1}) + V_{t+1}^\pi(S_{t+1}^\pi)) \text{ for all } s \in \mathcal{S}, t \in \mathcal{T}, \\ V_T^\pi(s) &= d_{T,\beta}(r, p) \text{ for all } s \in \mathcal{S}. \end{aligned}$$

We find that simply comparing  $V_0^\pi(S_0)$  to  $V_{0,\beta}^*(S_0)$  is not particularly illuminating due to the large “constant” introduced by  $f_t$ ; in fact, doing so results in our approximations



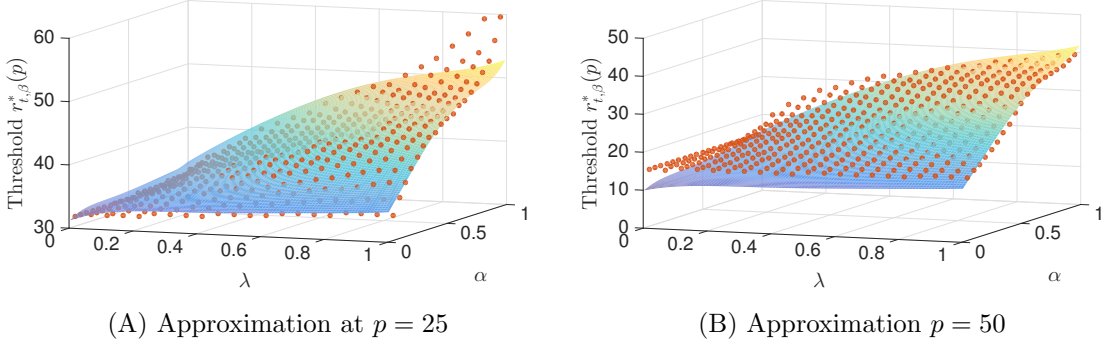


Figure 5: Illustration of Polynomial Approximation of  $r_{t,\beta}^*(p)$  for  $t = 12$

consistently achieving greater than 99% optimality. Instead, let  $\pi_d$  represent the default policy that charges the EV nonstop, the standard operation of an EV charging station; this is equivalent to setting the threshold at  $R_{\max}$  for any state where  $A_t = 1$ . We define:

$$\% \text{ optimality (with respect to default policy)} = \frac{V_0^\pi(S_0) - V_{0,\beta}^*(S_0)}{V_0^{\pi_d}(S_0) - V_{0,\beta}^*(S_0)}.$$

Table 1 tabulates the results after computing the percentage of optimality for the approximate polynomial policies  $\bar{r}_{t,\beta}(p)$ . We examine risk parameters at  $\lambda \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$  and  $\alpha \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ , which are points that are not used in the fitting procedure. The results are encouraging and suggest that the approximation has adequately captured the behaviors of the risk-averse optimal policies. However, we also note that they can be improved by increasing the degree of the polynomial approximation at a (modest) increase in CPU time.

$\beta = (\lambda, \alpha)$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.5$	$\alpha = 0.7$	$\alpha = 0.9$
$\lambda = 0.1$	93.8%	94.4%	94.9%	95.0%	95.5%
$\lambda = 0.3$	95.1%	96.7%	97.4%	97.8%	98.3%
$\lambda = 0.5$	96.2%	98.0%	98.7%	99.1%	99.4%
$\lambda = 0.7$	96.9%	98.6%	99.2%	99.5%	99.5%
$\lambda = 0.9$	97.1%	98.4%	99.0%	99.4%	98.8%

Table 1: Optimality of Approximate Policies  $\bar{r}_{t,\beta}(p)$

## 5.4 Approximate Pareto Frontier

Relevant questions regarding risk that a firm owning the EV charging stations may ask are: (1) what are the characteristics of the tradeoff between the cost of supplying energy,  $\text{Reward}(\pi_\beta^*)$ , versus the probability that a customer’s vehicle is inadequately charged,  $\text{Risk}^F(\pi_\beta^*; \delta)$ ? or (2) what are the characteristics of the tradeoff between the cost of supplying energy,  $\text{Reward}(\pi_\beta^*)$ , versus the cost of lost customers,  $\text{Risk}^D(\pi_\beta^*)$ ? These questions

can be investigated using the approximation  $\bar{\mathcal{R}}(\mathcal{B})$ .

Let  $\bar{\pi}_\beta$  refer to the policy at risk level  $\beta$  generated from  $\bar{\mathcal{R}}(\mathcal{B})$ . Using simulation and  $\bar{\mathcal{R}}(\mathcal{B})$ , we compute the metrics  $\text{Reward}(\bar{\pi}_\beta)$ ,  $\text{Risk}^F(\bar{\pi}_\beta; 0.2)$ , and  $\text{Risk}^D(\bar{\pi}_\beta)$  for a fine grid of  $\lambda$  and  $\alpha$  spaced at 0.005. A scatter plot of the risk/reward metrics for the roughly 40,000 policies is shown in Figure 6. The Pareto frontiers reside at the upper left borders of the shapes shown in Figure 6. Typically, we expect such a scatter plot to produce larger variation of points that covers a wider area. The “tightness” of our plots says that any policy generated from  $\bar{\mathcal{R}}(\mathcal{B})$  is *nearly* Pareto efficient for both types of risk/reward tradeoffs. In other words, for any  $\beta$ , the policy  $\bar{\pi}_\beta$  is such that we cannot significantly improve in one metric (i.e., either increase reward or decrease risk) without sacrificing in the other (i.e., either decrease reward or increase risk).

Therefore, the implication for managers of EV charging stations is that after obtaining  $\bar{\mathcal{R}}(\mathcal{B})$ , the decision of which  $\beta$  to choose need not be difficult: she simply needs find a policy  $\bar{\pi}_\beta$  that matches the firm’s risk or reward target (whichever is highest priority) and can be reasonably satisfied that  $\bar{\pi}_\beta$  is nearly Pareto efficient. This statement is purely an empirical observation for this specific problem, and whether or not it is a general phenomenon certainly merits further study. Finally, it is interesting to note that 47% of the time, the

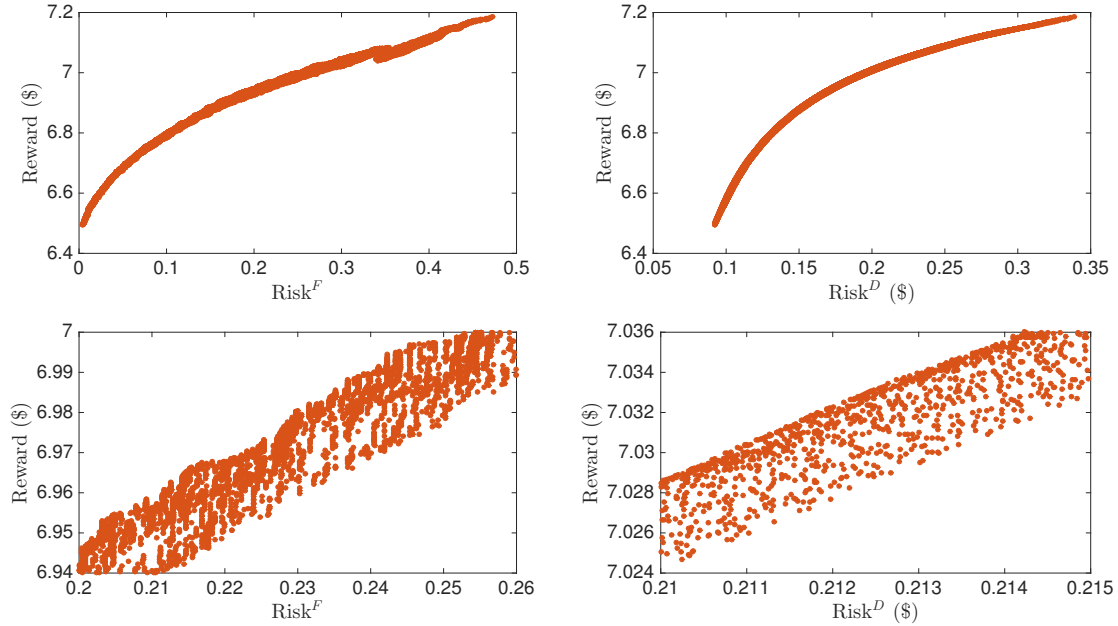


Figure 6: Reward vs. Risk Scatter Plots

optimal risk-neutral policy allows EVs to leave with less than 80% charge. The plot shows that we can find a  $\beta$  such that  $\text{Risk}^F(\bar{\pi}_\beta; 0.2)$  is reduced to nearly 0% if we are willing to forego 10% of the revenue (a reasonable tradeoff). A more general takeaway from this study is that our proposed approximation methodology allows for firms to easily conduct risk-based analyses (via simulation) on policies within the set  $\mathcal{R}^*(\mathcal{B})$ .

## 6 Conclusion

In this paper, we study EV charging in a setting where one attempts to exploit volatility in the spot market by charging only when the prices are low enough. At the same time, the risk of unsatisfied customers who may return to undercharged vehicles must be controlled. The problem is specified as a risk-averse MDP formulated using a dynamic risk measure, and the main result of the paper is a novel analysis of the behavior of the optimal policy with respect to changing the parameters  $\beta$  of the dynamic risk measure. We then propose using polynomial optimization and shape constraints to approximate the entire *spectrum of risk-averse policies*, i.e., all optimal risk-averse policies that can be generated by varying  $\beta$ . Next, we advocate that practitioners search through this set of policies, which have theoretical justification, in order to find a satisfactory operating policy. Finally, we provide numerical results that support the effectiveness of the approximation procedure and also show that various tradeoffs and questions of Pareto optimality can be easily investigated using the approximate spectrum of policies.

## A Proofs

**Proposition 1** (Convexity). *Let  $s = (r, p, 1) \in \mathcal{S}$  be the state variable. The following statements hold for every  $t$  and  $\beta_t$ : (i)  $xp + \tilde{V}_{t,\beta}(r+x, p)$  is strictly convex in  $x$  on  $\mathcal{X}(r)$ , (ii)  $\tilde{V}_{t,\beta}(s)$  is strictly convex in the resource  $r$ , and (iii)  $V_{t,\beta}^*(s)$  is strictly convex in the resource  $r$ .*

*Proof.* First note that  $V_{t,\beta}^*(r, p, 0)$  is strictly convex in  $r$  for every  $t$  by (12). Properties (i) and (ii) hold for the base case  $t = T$  by (12), and let the induction hypothesis be that  $V_{t+1,\beta}^*(s)$  is strictly convex in  $r$ . For  $x \in \mathcal{X}(r)$ , let

$$Q_{t,\beta}^*(s, x) = xp + \tilde{V}_{t,\beta}(r+x, p)$$

and  $\mathcal{Q} = \{(r, x) : x \in \mathcal{X}(r), r \in \mathcal{R}\}$ . Using the convexity (in the space of random variables) of  $\rho_{\beta_t}$  and the induction hypothesis, we can easily verify (ii). Similarly, using the strict convexity of the mapping  $(r, x) \mapsto V_{t+1,\beta}^*(r+x, \hat{Z}_{t+1}(s))$  on  $\mathcal{Q}$  for each realization of  $\hat{Z}_{t+1}(s)$ , we obtain that

$$(r, x) \mapsto \tilde{V}_{t,\beta}(r+x, p) = \rho_{\beta_t}[V_{t+1,\beta}^*(r+x, \hat{Z}_{t+1}(s))]$$

is strictly convex on  $\mathcal{Q}$ , from which we see that  $Q_{t,\beta}^*(s, x)$  is strictly convex in  $(r, x)$  on  $\mathcal{Q}$ . Since we can write  $V_{t,\beta}^*(s) = \min_{x \in \mathcal{X}(r)} Q_{t,\beta}^*(s, x)$  and  $\mathcal{X}(r)$  is nonempty, the well-known property that convexity is preserved under minimization can be applied; by [Heyman and Sobel, 2003, Proposition B-4],  $V_{t,\beta}^*(s)$  is strictly convex in  $r$ . We have completed the inductive step and also verified (i) and (ii) along the way.  $\square$

**Lemma 1** (Lipschitz Property). *For any  $r, r' \in \mathcal{R}$ ,  $p, p' \in \mathbb{R}$ , and  $a \in \{0, \pm 1\}$  and for each  $t \leq T$ , there exists deterministic  $L_t > 0$  such that the optimal value function  $V_{t,\beta}^*$  satisfies*

$$|V_{t,\beta}^*(r, p, a) - V_{t,\beta}^*(r', p', a)| \leq L_t \|(r, p) - (r', p')\|_2,$$

where  $\|\cdot\|_2$  is the Euclidean norm.

*Proof.* We only need to consider  $a = 1$  and proceed via backward induction on  $t$ . The statement of the lemma is clearly true for the base case,  $t = T$ , so we assume that it holds for  $t+1$ . Notice that we can equivalently write  $V_{t,\beta}^*(r, p, a) = \min_{x \in \mathcal{X}} xp + \tilde{V}_{t,\beta}(r+x \wedge R_{\max}, p)$ , as it is never optimal to choose  $x$  such that  $r+x > R_{\max}$ . Therefore, using the inequality  $|\min f - \min g| \leq \max |f - g|$ ,

$$\begin{aligned} & |V_{t,\beta}^*(r, p, a) - V_{t,\beta}^*(r', p', a)| \\ & \leq \max_{x \in \mathcal{X}} |\tilde{V}_{t,\beta}(r+x \wedge R_{\max}, p) - \tilde{V}_{t,\beta}(r'+x \wedge R_{\max}, p')| + x_{\max} \|(r, p) - (r', p')\|_2. \end{aligned}$$

Recall from (7) that since  $P_t$  is an OU process,  $P_{t+1} | P_t = p$  is equal in distribution to  $\mu_P(p) + \epsilon_{v_P}$  and  $P_{t+1} | P_t = p'$  is equal in distribution to  $\mu_P(p') + \epsilon_{v_P}$ , where  $\epsilon_{v_P}$  is  $\mathcal{N}(0, v_P)$ .

Thus, the law invariance of  $\rho_{\beta_t}$  gives

$$\tilde{V}_{t,\beta}(r, p) = \rho_{\beta_t} [V_{t+1,\beta}^*(r, \mu_P(p) + \epsilon_v, A_{t+1}) | A_t = 1].$$

Now using  $|\rho_{\beta_t}(A) - \rho_{\beta_t}(B)| \leq \rho_{\beta_t}(|A - B|)$  (which follows by the convexity of  $\rho_{\beta_t}$ ), the induction hypothesis, and the fact that  $(r + x \wedge R_{\max}) - (r' + x \wedge R_{\max}) \leq |r - r'|$  we obtain

$$|V_{t,\beta}^*(r, p, a) - V_{t,\beta}^*(r', p', a)| \leq L_{t+1} \|(r, \mu_P(p)) - (r', \mu_P(p'))\|_2 + x_{\max} \|(r, p) - (r', p')\|_2.$$

We can easily conclude by noting that  $\mu$  is Lipschitz with constant  $e^{-\kappa_P}$ .  $\square$

**Proposition 2** (Monotonicity in Spot Price). *Suppose the conditions of Assumption 1 are satisfied. For  $t \leq T$ ,  $r \in \mathcal{R}$ , and  $a \in \{0, 1\}$ , the optimal value function  $V_{t,\beta}^*(r, p, a)$  is (strictly) increasing in the spot price  $p$ .*

*Proof.* Again, we only need to consider  $a = 1$ , so we let  $s = (r, p, 1)$ . By (12), for each  $r$ ,  $V_{T,\beta}^*(r, p, 1) = d_{T,\beta}(r, p)$  is a decreasing function in the spot price  $p$  by Assumption 1. This completes the base case. We assume the same is true for  $V_{t+1,\beta}^*(r, p, 1)$  and aim to show that  $xp + \tilde{V}_{t,\beta}(r + x, p)$  is increasing in  $p$  for fixed resource  $r \in \mathcal{R}$  and decision  $x \in \mathcal{X}(r)$ . Note that  $xp$  is certainly nonincreasing in  $p$  but it is not strictly monotone for  $x = 0$ , which leaves us to show  $\tilde{V}_{t,\beta}(r + x, p)$  is increasing in  $p$ . Define

$$g(p) = \mathbf{E}[V_{t+1,\beta}^*(r + x, \hat{Z}_{t+1}(s))],$$

so that we have  $\tilde{V}_{t,\beta}(r + x, p) = (1 - \lambda_t)g(p) + \lambda_t c_{t+1}^*(r + x, p)$ . Since  $\lambda_t$  can take values of 0 or 1, we must show that both  $g(\cdot)$  and  $c_{t+1}^*(r + x, \cdot)$  are increasing functions. As before,  $P_{t+1} | P_t = p$  is equal in distribution to  $\mu_P(p) + \epsilon_{v_P}$ , where  $\epsilon_{v_P} \sim \mathcal{N}(0, v_P)$ . Let  $q_t^0 = \mathbf{P}(A_{t+1} = 0 | A_t = 1)$  and  $q_t^1 = \mathbf{P}(A_{t+1} = 1 | A_t = 1)$ . Using the independence of  $\tau$  and the price process  $\{P_t\}$ , we have

$$g(p) = q_t^0 \mathbf{E}[d_{t+1,\beta}(r + x, \mu_P(p) + \epsilon_{v_P})] + q_t^1 \mathbf{E}[V_{t+1,\beta}^*(r + x, \mu_P(p) + \epsilon_{v_P}, 1)].$$

By the induction hypothesis, the terms within the expectation are increasing in  $p$  for every realization of  $\epsilon_v$ , so  $f_e(p)$  is increasing in  $p$ . Now, we consider  $c_{t+1}^*(r + x, p)$ . By the representation given in (10), we can write CVaR in terms of VaR as follows:

$$c_{t,\gamma}^*(r + x, p) = \frac{1}{1 - \gamma} \int_{\gamma}^1 v_{t,w}^*(r + x, p) dw.$$

By definition, the VaR is given by the equation

$$v_{t,w}^*(r + x, p) = \inf \left\{ u : \sum_{a \in \{0,1\}} q_t^a \mathbf{P}(V_{t+1,\beta}^*(r + x, \mu_P(p) + \epsilon_{v_P}, a) \leq u) > w \right\},$$

from which it is clear that  $v_{t,w}^*(r + x, p)$  is increasing in  $p$ . It thus follows that  $c_t^*(r + x, p)$  is increasing in the spot price  $p$ .  $\square$

**Lemma 3** (Differentiability of Value Function in Spot Price). *Suppose that for some  $t$  and resource  $r$ , the partial derivative  $\partial_p \tilde{V}_{t,\beta}(r, p)$  exists for any spot price  $p$  and is nonzero. Then, it follows that  $\partial_p V_{t,\beta}^*(r, p, 1)$  also exists for any spot price  $p$  and is nonzero.*

*Proof.* Let  $s = (r, p, 1)$ . The case of  $r = R_{\max}$  is easy because the optimal decision is  $x_{t,\beta}^*(R_{\max}, p) = 0$ . For  $r < R_{\max}$ , Slater's condition holds and we can apply the envelope theorem (with inequality constraints) to (11), obtaining

$$\partial_p V_{t,\beta}^*(r, p, 1) = x_{t,\beta}^*(r, p) + \partial_p \tilde{V}_{t,\beta}(r + x, p)|_{x=x_{t,\beta}^*(r,p)}, \quad (21)$$

so the conclusion of (i) follows because  $x_{t,\beta}^*(r, p) \geq 0$  and  $\partial_p \tilde{V}_{t,\beta}(r + x, p)|_{x=x_{t,\beta}^*(r,p)} > 0$  (see the proof of Proposition 2).  $\square$

**Lemma 4** (Interchanging Derivative and Risk Measure). *We consider differentiation in  $r$  and  $p$  separately, starting with  $r$ . Fix a spot price  $p$  and let  $s = (r, p, 1)$ . Let  $\bar{\mathcal{R}} \subseteq \mathcal{R}$  be a set of resource values where  $V_{t+1,\beta}^*(r, p, 1)$  is differentiable. Then, for every  $r \in \bar{\mathcal{R}}$ ,*

$$\begin{aligned} \partial_r \tilde{V}_{t,\beta}(r, p) &= (1 - \lambda_t) \mathbf{E}[\partial_r V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s))] \\ &\quad + \lambda_t \mathbf{E}[\partial_r V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s)) \mid V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s)) \geq v_{t,\alpha}^*(r, p)]. \end{aligned}$$

Now, we fix a resource state  $r$ . Suppose that  $V_{t+1,\beta}^*(r, p, 1)$  is differentiable everywhere in the spot price. It follows that for every  $p \in \mathbb{R}$ ,

$$\begin{aligned} \partial_p \tilde{V}_{t,\beta}(r, p) &= (1 - \lambda_t) \mathbf{E}[\partial_p V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s))] \\ &\quad + \lambda_t \mathbf{E}[\partial_p V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s)) \mid V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s)) \geq v_{t,\alpha}^*(r, p)]. \end{aligned}$$

We remark that “interchange” is used loosely here; the second term (in both cases above) resembles the “CVaR of the derivative” but note that the conditioning event has not changed.

*Proof.* The arguments for  $\partial_r \tilde{V}_{t,\beta}(r, p)$  and  $\partial_p \tilde{V}_{t,\beta}(r, p)$  are nearly identical. In both cases, the first part of the equation is simply an interchange of differentiation and expectation, given by the equations

$$\begin{aligned} \partial_r \mathbf{E}[V_{t+1}^*(r, \hat{Z}_{t+1}(s))] &= \mathbf{E}[\partial_r V_{t+1}^*(r, \hat{Z}_{t+1}(s))], \\ \partial_p \mathbf{E}[V_{t+1}^*(r, \hat{Z}_{t+1}(s))] &= \mathbf{E}[\partial_p V_{t+1}^*(r, \hat{Z}_{t+1}(s))]. \end{aligned}$$

The justification for both cases is due to Lemma 1 and the dominated convergence theorem. The second term follows by an application of [Hong and Liu, 2009, Theorem 3.1]; however, we must verify the three assumptions of the theorem, [Hong and Liu, 2009, Assumptions 1–3]. The *first assumption* states that the function must be Lipschitz in the parameter (either  $r$  or  $p$  in this case) and that the derivative in the parameter exists almost surely. The Lipschitz property is readily verified by Lemma 1. The almost sure existence of  $\partial_r V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s))$  is verified by keeping in mind that  $V_{t+1,\beta}^*(r, p, 0) = d_{t+1,\beta}^*(r, p)$ , which is differentiable by Assumption 1. The almost sure existence of  $\partial_p V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s))$  follows

because  $P_{t+1} | P_t = p$  can be written as  $\mu_P(p) + \epsilon_{v_P}$ . The *second assumption* requires the value at risk function  $v_{t,\alpha}^*(r, p)$  to be differentiable in the parameter. When  $r$  is the parameter, we notice that by Lemma 2, the partial derivative  $\partial_r v_{t,\alpha}^*(r, p)$  exists whenever  $r \in \bar{\mathcal{R}}$ . When  $p$  is the parameter,  $\partial_p v_{t,\alpha}^*(r, p)$  always exists by Lemma 2. The *third assumption* states that  $\{V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s)) = v_{t,\alpha}^*(r, p)\}$  is a zero probability event. It is clear that by Proposition 2, we have for any  $v$ ,

$$\mathbf{P}(V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s)) = v) = \sum_{a \in \{0,1\}} q_t^a \mathbf{P}(\epsilon_{v_P} = \nu_{t+1}^a(r, v) - \mu_P(p)) = 0.$$

Hence, the assumptions are verified and we conclude using the sensitivity result of [Hong and Liu, 2009, Theorem 3.1].  $\square$

**Theorem 3** (Threshold Behavior). *Under Assumptions 1–2, the following statements regarding the behavior of the optimal threshold  $r_{t,\beta}^*(p)$  hold.*

- (i) *The threshold  $r_{t,\beta}^*(p)$  is nondecreasing in the risk parameters  $\beta_t = (\lambda_t, \alpha_t)$ . In other words, as risk aversion increases, the thresholds relax.*
- (ii) *The threshold  $r_{t,\beta}^*(p)$  is nonincreasing in the spot price  $p$ .*

*Proof.* The proof is by backward induction on  $t$ . We employ an induction hypothesis with multiple parts, enumerated below:

- (a) For every spot price  $p$ , the value function  $\tilde{V}_{t,\beta}$  is piecewise continuously differentiable.
- (b) For every resource state  $r$ , the optimal value function  $\tilde{V}_{t,\beta}$  is differentiable in  $p$  at every spot price  $p$  and the derivative is nonzero.
- (c) The quantity  $p + \mathbf{E}[\partial_r V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s)) | V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s)) \geq v_{t,\alpha}^*(r, p)]$  is nondecreasing in the spot price  $p$  whenever the derivative in  $r$  exists.
- (d) The quantity  $p + \tilde{V}'_{t,\beta}(r, p)$  is nondecreasing in the spot price  $p$ . This implies part (ii) of the theorem.
- (e) The derivative of the value function  $\tilde{V}'_{t,\beta}(r, p)$  is nonincreasing in the spot price  $p$ .
- (f) The derivative of the value function  $\tilde{V}'_{t,\beta}(r, p)$  is componentwise nonincreasing in  $\beta_t = (\lambda_t, \alpha_t)$ . This implies part (i) of the theorem.

Let us consider the base case,  $t + 1 = T$ . By (12) and Assumption 1, we know that  $V_{T,\beta}^*(r, p) = d_{T,\beta}(r, p)$  is continuously differentiable in  $r$ . By Lemma 4 and dominated convergence (the derivatives are bounded by part (iii) of Assumption 1), we know that  $\tilde{V}_{T-1,\beta}(r, p)$  is continuously differentiable and thus, trivially piecewise continuously differentiable. Part (b) holds via a similar argument. Part (c) holds because  $V_{t+1,\beta}^*(r, p, 1) = V_{t+1,\beta}^*(r, p, 0) = d_{t+1,\beta}(r, p)$  and part (ii) Assumption 1, which states that the derivatives



in the resource are unaffected by the spot price  $p$ . Parts (d), (e), and (f) hold because of Lemma 4, also in conjunction with part (ii) of Assumption 1. We assume the induction hypothesis (referred to as the “ $t+1$  induction hypothesis”) and move on to the inductive step, where we aim to show parts (a) through (e) of the induction hypothesis with  $t$  replacing  $t+1$  (referred to as the “ $t$  induction hypothesis”).

By Assumption 2 and part (a) of the induction hypothesis,  $V_{t,\beta}^*(r, p, 1)$  piecewise continuously differentiable, from which it readily follows by Lemma 4 and the dominated convergence theorem that  $\tilde{V}_{t-1,\beta}(r, p)$  is also piecewise continuously differentiable, so part (a) of the inductive step is complete. Part (b) of the inductive step is true via a similar analysis, but applying Lemma 3 instead of Assumption 2.

For the remaining parts, (c), (d), (e), and (f), we first consider differentiable resource states; fix a resource  $r$  at which  $\tilde{V}_{t-1,\beta}$  and  $V_{t,\beta}^*$  are differentiable and move on to part (c). We now use  $P$  to represent an outcome of the spot price at time  $t$  in order to distinguish from  $p$ , the spot price at  $t-1$ . Recall that the optimal policy at time  $t$ , by Theorem 2, is to charge up to (or as close as possible to) the threshold  $r_{t,\beta}^*(P)$ . We define the sets  $\mathcal{P}_{\text{low}} = \{P : r + x_{\max} \leq r_{t,\beta}^*(P)\}$ ,  $\mathcal{P}_{\text{med}} = \{P : r \leq r_{t,\beta}^*(P) < r + x_{\max}\}$ , and  $\mathcal{P}_{\text{high}} = \{P : r_{t,\beta}^*(P) \leq r\}$ , where it is possible for some of these sets to be empty. The threshold  $r_{t,\beta}^*(P)$  is continuous in  $P$  by Theorem 2 and nonincreasing in  $P$  by the  $t+1$  induction hypothesis (c). Hence, we can find breakpoints  $a_{t,\beta}(r), b_{t,\beta}(r) \in \mathbb{R} \cup \{\pm\infty\}$  to equivalently write the above sets as intervals

$$\mathcal{P}_{\text{low}} = (-\infty, a_{t,\beta}(r)], \mathcal{P}_{\text{med}} = (a_{t,\beta}(r), b_{t,\beta}(r)], \text{ and } \mathcal{P}_{\text{high}} = (b_{t,\beta}(r), \infty).$$

By the continuity of the threshold  $r_{t,\beta}^*(p)$  in  $p$ , we suppose that when  $a_{t,\beta}(r)$  is finite, it is chosen so that when  $p = a_{t,\beta}(r)$ , it holds that  $r_{t,\beta}^*(p) = r + x_{\max}$ . Similarly, when  $b_{t,\beta}(r)$  is finite,  $p = b_{t,\beta}(r)$  implies  $r_{t,\beta}^*(p) = r$ . Doing casework on each of these spot price intervals, we have

$$V_{t,\beta}^*(r, P, 1) = \begin{cases} h(P, x_{\max}) + \tilde{V}_{t,\beta}(r + x_{\max}, P) & \text{if } P \in \mathcal{P}_{\text{low}}, \\ h(P, r_{t,\beta}^*(P) - r) + \tilde{V}_{t,\beta}(r_{t,\beta}^*(P), P) & \text{if } P \in \mathcal{P}_{\text{med}}, \\ h(P, 0) + \tilde{V}_{t,\beta}(r, P) & \text{if } P \in \mathcal{P}_{\text{high}}. \end{cases}$$

Taking derivatives and adding  $P$  to both sides, we obtain

$$w_{t,\beta}(r, P) := P + \partial_r V_{t,\beta}^*(r, P, 1) = \begin{cases} P + \tilde{V}'_{t,\beta}(r + x_{\max}, P) & \text{if } P \in \mathcal{P}_{\text{low}}, \\ 0 & \text{if } P \in \mathcal{P}_{\text{med}}, \\ P + \tilde{V}'_{t,\beta}(r, P) & \text{if } P \in \mathcal{P}_{\text{high}}. \end{cases}$$

By the  $t+1$  induction hypothesis part (d),  $w_{t,\beta}(r, P)$  is “piecewise nondecreasing” (i.e., nondecreasing when restricted to each of the three intervals) in  $P$ . However, we must check that it does not decrease at the two boundaries between the intervals. Recall the

optimization problem for computing the threshold  $r_{t,\beta}^*(P)$ :

$$\min_{\tilde{r} \in \mathcal{R}} \tilde{r} P + \tilde{V}_{t,\beta}(\tilde{r}, P), \quad (22)$$

and notice that the derivative of the objective in  $\tilde{r}$  is given by  $P + \tilde{V}'_{t,\beta}(\tilde{r}, P)$ . By convexity and the fact that  $r + x_{\max}$  is to the left of the optimal solution, we can infer that for  $P \in \mathcal{P}_{\text{low}} = \{p : r + x_{\max} \leq r_{t,\beta}^*(p)\}$  it must be the case that  $P + \tilde{V}'_{t,\beta}(r + x_{\max}, P) \leq 0$ . Similarly,  $P + \tilde{V}'_{t,\beta}(r, P) \geq 0$  for  $P \in \mathcal{P}_{\text{high}}$ , which is enough to show that  $w_{t,\beta}(r, P)$  is nondecreasing in  $P$ .

We now show part (c) of the  $t$  induction hypothesis. Note that  $V_{t,\beta}^*(r, \hat{Z}_t(s))$  is a mixture distribution (depending on the arrival or nonarrival of the customer) created from  $V_{t,\beta}^*(r, P_t(p), 0)$  and  $V_{t,\beta}^*(r, P_t(p), 1)$ , whose distribution functions are denoted  $F_t^0$  and  $F_t^1$ , respectively. It turns out that it is useful for us to be able to refer to the risk level at which  $v_{t-1,\alpha}^*(r, p)$  falls for each of the component distributions. Hence, we define the mappings:

$$\Lambda_t^0(\alpha, p) = F_t^0(v_{t-1,\alpha}^*(r, p)) \quad \text{and} \quad \Lambda_t^1(\alpha, p) = F_t^1(v_{t-1,\alpha}^*(r, p)),$$

both of which are nondecreasing in  $p$  and  $\alpha$ . For both  $a \in \{0, 1\}$ , it thus follows that the VaR at level  $\Lambda_t^a(p, \alpha)$  of  $V_{t,\beta}^*(r, P_t(p), 0)$  is equal to  $v_{t-1,\alpha}^*(r, p)$ . Now let us consider the events  $\{V_{t,\beta}^*(r, \hat{Z}_t(s)) \geq v_{t-1,\alpha}^*(r, p), A_t = a\}$  for  $a \in \{0, 1\}$ :

$$\begin{aligned} & \{V_{t,\beta}^*(r, \hat{Z}_t(s)) \geq v_{t-1,\alpha}^*(r, p), A_t = a\} \\ &= \{V_{t,\beta}^*(r, P_t(p), a) \geq \text{VaR}_{\Lambda_t^a(\alpha, p)}[V_{t,\beta}^*(r, P_t(p), a)], A_t = a\} \\ &= \{V_{t,\beta}^*(r, P_t(p), a) \geq V_{t,\beta}^*(r, \text{VaR}_{\Lambda_t^a(\alpha, p)}(P_t(p)), a), A_t = a\} \\ &= \{P_t(p) \geq \text{VaR}_{\Lambda_t^a(\alpha, p)}(P_t(p)), A_t = a\}, \end{aligned} \quad (23)$$

where we have applied Proposition 2 and the property that VaR is invariant under monotone transformations. Therefore, by independence of  $\{A_t\}$  and  $\{P_t\}$ , the quantity in question for part (c) of the  $t$  induction hypothesis is

$$p + \sum_{a \in \{0,1\}} q_{t-1}^a \mathbf{E}[\partial_r V_{t,\beta}^*(r, P_t(p), a) \mid P_t(p) \geq \text{VaR}_{\Lambda_t^a(\alpha, p)}(P_t(p))]. \quad (24)$$

Since  $\partial_r V_{t,\beta}^*(r, P_t(p), 0) = \partial_r d_{t,\beta}(r, P_t(p))$ , the term for  $a = 0$  is trivially nondecreasing in  $p$  by part (ii) of Assumption 1. Rewriting  $P_t(p)$  as  $\mu_P(p) + \epsilon_{v_P}$ , the remaining term can be simplified:

$$\begin{aligned} & p + q_{t-1}^1 \mathbf{E}[\partial_r V_{t,\beta}^*(r, P_t(p), 1) \mid P_t(p) \geq \text{VaR}_{\Lambda_t^1(\alpha, p)}(P_t(p))] \\ &= p + q_{t-1}^1 \mathbf{E}[w_{t,\beta}(r, P_t(p)) - P_t(p) \mid P_t(p) \geq \text{VaR}_{\Lambda_t^1(\alpha, p)}(P_t(p))] \\ &= (p - q_{t-1}^1 \mu_P(p)) + q_{t-1}^1 \mathbf{E}[w_{t,\beta}(r, \mu_P(p) + \epsilon_{v_P}) \mid \epsilon_{v_P} \geq \text{VaR}_{\Lambda_t^1(\alpha, p)}(\epsilon_{v_P})]. \end{aligned}$$

Using  $q_{t-1}^1 e^{-\kappa_P} < 1$  and the nondecreasing nature of  $w_{t+1,\beta}(r, P)$  and  $\Lambda_t^1(p, \alpha)$ , it is clear that the entire term is nondecreasing in  $p$ . Part (c) for differentiable  $r$  is complete.

Moving on to part (d), the next equation follows by Lemma 4:

$$p + \partial_r \tilde{V}_{t-1,\beta}(r, p) = (1 - \lambda_{t-1}) \left[ p + \mathbf{E}[\partial_r V_{t,\beta}^*(r, \hat{Z}_{t+1}(s))] \right] + \lambda_{t-1} \left[ p + \mathbf{E}[\partial_r V_{t,\beta}^*(r, \hat{Z}_t(s)) \mid V_{t,\beta}^*(r, \hat{Z}_t(s)) \geq v_{t-1,\alpha}^*(r, p)] \right]. \quad (25)$$

The first term in brackets can be written as

$$\begin{aligned} p + \mathbf{E}[\partial_r V_{t,\beta}^*(r, \hat{Z}_t(s))] &= p + \sum_{a \in \{0,1\}} q_{t-1}^a \mathbf{E}[\partial_r V_{t,\beta}^*(r, P_t(p), a)] \\ &= p + q_{t-1}^0 \mathbf{E}[\partial_r d_{t,\beta}(r, P_t(p))] + q_{t-1}^1 \mathbf{E}[w_{t,\beta}(r, P_t(p)) - P_t(p)] \\ &= [p - q_{t-1}^1 \mu_P(p)] + q_{t-1}^0 \mathbf{E}[\partial_r d_{t,\beta}(r, P_t(p))] + q_{t-1}^1 \mathbf{E}[w_{t,\beta}(r, P_t(p))]. \end{aligned}$$

The first term  $p - q_{t-1}^1 \mu_P(p)$  is increasing in  $p$  because  $q_{t-1}^1 e^{-\kappa_P} < 1$ , and the second term  $\mathbf{E}[\partial_r d_{t,\beta}(r, P_t(p))]$  is independent of  $p$  given Assumption 1. Replacing  $P_t(p)$  with  $\mu_P(p) + \epsilon_{v_P}$  and applying the nondecreasing property of  $w_{t,\beta}(r, P)$ , it is clear that the final term is nondecreasing in  $p$ , so we have confirmed that  $p + \mathbf{E}[\partial_r V_{t+1,\beta}^*(r, \hat{Z}_{t+1}(s))]$  is nondecreasing in  $p$ . The second term of (25), given by  $p + \mathbf{E}[\partial_r V_{t,\beta}^*(r, \hat{Z}_t(s)) \mid V_{t,\beta}^*(r, \hat{Z}_t(s)) \geq v_{t-1,\alpha}^*(r, p)]$ , is nondecreasing by part (c) of the  $t$  induction hypothesis, which is proven above. We conclude that  $p + \partial_r \tilde{V}_{t,\beta}(r, p)$  is nondecreasing in  $p$ , as required for part (d) of the inductive step.

Let us now prove part (e) of the  $t$  induction hypothesis. Using the same steps as before, we have

$$\partial_r V_{t,\beta}^*(r, P, 1) = \begin{cases} \tilde{V}'_{t,\beta}(r + x_{\max}, P) & \text{if } P \in \mathcal{P}_{\text{low}}, \\ -P & \text{if } P \in \mathcal{P}_{\text{med}}, \\ \tilde{V}'_{t,\beta}(r, P) & \text{if } P \in \mathcal{P}_{\text{high}}, \end{cases}$$

which, by part (e) of the  $t+1$  inductive hypothesis, is “piecewise nonincreasing.” We aim to show that  $\partial_r V_{t,\beta}^*(r, P, 1)$  is nonincreasing in  $P$ , so it remains for us to check that the function does not decrease (i.e., jump) at the either of the two boundaries between the intervals. Consider  $a_{t,\beta}(r)$ , the boundary between  $\mathcal{P}_{\text{low}}$  and  $\mathcal{P}_{\text{med}}$  and assume that this boundary exists (if it does not, then either  $\mathcal{P}_{\text{low}} = \mathbb{R}$  or  $\mathcal{P}_{\text{low}} = \emptyset$  and we can move on). Similar to before and using our choice of  $a_{t,\beta}(r)$ , when  $P \leq a_{t,\beta}(r)$  or  $r + x_{\max} \leq r_{t,\beta}^*(p)$ , the point  $r + x_{\max}$  is to the left of the optimal solution  $r_{t,\beta}^*(p)$  and so it follows from (22) that the derivative  $P + \tilde{V}'_{t,\beta}(r + x_{\max}, P) \leq 0$ . On the other hand, when  $b_{t,\beta}(r) \geq P \geq a_{t,\beta}(r)$  or  $r \leq r_{t,\beta}^*(P) \leq r + x_{\max}$ , we know that  $r$  is still to the left of  $r_{t,\beta}^*(P)$ , but it must be true that  $r + x_{\max}$  is to the right of  $r_{t,\beta}^*(P)$ . Hence,  $P + \tilde{V}'_{t,\beta}(r + x_{\max}, P) \geq 0$ . Since the boundary point  $a_{t,\beta}(r)$  is included in both intervals (this works because  $r_{t,\beta}^*(a_{t,\beta}(r)) = r + x_{\max}$ ), we see that  $-P = \tilde{V}'_{t,\beta}(r + x_{\max}, P)$ . Therefore,  $\partial_r V_{t,\beta}^*(r, P, 1)$  is nonincreasing as  $P$  moves from  $\mathcal{P}_{\text{low}}$  to  $\mathcal{P}_{\text{med}}$ . A completely analogous argument can be used to show that the same holds when  $P$  moves from  $\mathcal{P}_{\text{med}}$  to  $\mathcal{P}_{\text{high}}$ . Therefore,  $\partial_r V_{t,\beta}^*(r, P, 1)$  is nonincreasing in  $P$ . Applying

Lemma 4 and performing a similar analysis as in part (d), we conclude that  $\partial_r \tilde{V}_{t-1}(r, p)$  is nonincreasing in the spot price  $p$ .

For part (f), we first consider  $\alpha_{t-1}$ . By Lemma 4 and (23), we need to show that the following quantity is nonincreasing in the risk parameter  $\alpha_{t-1}$ :

$$\begin{aligned} \mathbf{E}[\partial_r V_{t,\beta}^*(r, \hat{Z}_t(s)) | V_{t,\beta}^*(r, \hat{Z}_t(s)) \geq v_{t-1,\alpha}^*(r, p)] \\ = \sum_{a \in \{0,1\}} q_{t-1}^a \mathbf{E}[\partial_r V_{t,\beta}^*(r, P_t(p), a) | P_t(p) \geq \text{VaR}_{\Lambda_t^a(\alpha, p)}(P_t(p))] \\ = \sum_{a \in \{0,1\}} q_{t-1}^a \mathbf{E}[\partial_r V_{t,\beta}^*(r, \mu_P(p) + \epsilon_{v_P}, a) | \epsilon_{v_P} \geq \text{VaR}_{\Lambda_t^a(\alpha, p)}(\epsilon_{v_P})]. \end{aligned}$$

By the above analysis within the proof of part (e) and recalling Assumption 1, we know that  $\partial_r V_{t,\beta}^*(r, P, a)$  is nonincreasing for both values of  $a$ . Because  $\Lambda_t^a(\alpha, p)$  is nondecreasing in  $\alpha_{t-1}$ , it follows that  $\text{VaR}_{\Lambda_t^a(\alpha, p)}(\epsilon_{v_P})$  is nondecreasing in  $\alpha_{t-1}$ . Combining these two results, we conclude that the expectation is nonincreasing in  $\alpha_{t-1}$ . For  $\lambda_{t-1}$ , note that since  $\partial_r V_{t,\beta}^*(r, P, a)$  is nonincreasing, conditioning on  $\{\epsilon_{v_P} \geq \text{VaR}_{\Lambda_t^a(\alpha, p)}(\epsilon_{v_P})\}$  can only decrease the value of the expectation; more precisely,

$$\begin{aligned} \mathbf{E}[\partial_r V_{t,\beta}^*(r, \hat{Z}_t(s))] &\geq \sum_{a \in \{0,1\}} q_{t-1}^a \mathbf{E}[\partial_r V_{t,\beta}^*(r, \mu_P(p) + \epsilon_{v_P}, a) | \epsilon_{v_P} \geq \text{VaR}_{\Lambda_t^a(\alpha, p)}(\epsilon_{v_P})], \\ &= \mathbf{E}[\partial_r V_{t,\beta}^*(r, \hat{Z}_t(s)) | V_{t,\beta}^*(r, \hat{Z}_t(s)) \geq v_{t-1,\alpha}^*(r, p)]. \end{aligned}$$

From Lemma 4, this implies that increasing  $\lambda_t$  cannot make  $\partial_r \tilde{V}_{t,\beta}(r, p)$  larger and part (f) of the inductive step is complete.

Now that we have shown the inductive step for the points of differentiability, we need to also consider the nondifferentiable points. Suppose  $\bar{r} \neq 0$  is a nondifferentiable point. Then, by part (a) of the inductive step for time  $t$ , it is true that the left derivative can be written as a limit of the derivatives of nearby points of differentiability, i.e.,  $\tilde{V}'_{t,\beta}(\bar{r}, p) = \lim_{r \uparrow \bar{r}} \partial_r \tilde{V}_{t,\beta}(r, p)$ . Therefore, the properties that hold for  $\partial_r \tilde{V}_{t,\beta}(r, p)$  carry over to the left derivative at  $\tilde{V}'_{t,\beta}(\bar{r}, p)$ . The proof is complete.  $\square$

## References

- C. Acerbi and D. Tasche. On the coherence of expected shortfall. *Journal of Banking and Finance*, 26(7):1487–1503, 2002.
- A. A. Ahmadi and A. Majumdar. DSOS and SDSOS optimization: LP and SOCP-based alternatives to sum of squares optimization. In *2014 IEEE Conference on Information Sciences and Systems (CISS)*, pages 1–5. IEEE, 2014.
- A. A. Ahmadi and A. Majumdar. Some applications of polynomial optimization in operations research and real-time decision making. *Optimization Letters*, pages 1–21, 2015.
- P. Artzner, F. Delbaen, J. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.
- P. Bansal. Charging of electric vehicles: Technology and policy implications. *Journal of Science Policy & Governance*, 6(1), 2015.
- S. Bashash, S. J. Moura, and H. K. Fathy. Charge trajectory optimization of plug-in hybrid electric vehicles for energy cost reduction and battery health enhancement. *American Control Conference (ACC), 2010*, pages 5824–5831, 2010.
- S. Bashash, S. J. Moura, J. C. Forman, and H. K. Fathy. Plug-in hybrid electric vehicle charge pattern optimization for energy cost and battery longevity. *Journal of Power Sources*, 196(1):541–549, 2011.
- Bay Area Council. Electrify your business: Moving forward with electric vehicles – A Bay area business guide. 2011.
- I. S. Bayram, G. Michailidis, M. Devetsikiotis, S. Bhattacharya, A. Chakraborty, and F. Granelli. Local energy storage sizing in plug-in hybrid electric vehicle charging stations under blocking probability constraints. *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 78–83, 2011.
- C. Berge. *Topological Spaces*. Oliver & Boyd, Edinburgh and London, 1963.
- K. C. Border. *Fixed point theorems with applications to economics and game theory*. Cambridge University Press, 1989.
- O. Çavus and A. Ruszczyński. Computational methods for risk-averse undiscounted transient markov models. *Operations Research*, 62(2):401–417, 2014.
- D. Chang, D. Erstad, E. Lin, A. F. Rice, C. T. Goh, A. A. Tsao, and J. Snyder. Financial viability of non-residential electric vehicle charging stations. *Luskin Center for Innovation: Los Angeles, CA, USA.*, 2012.

- P. Cheridito, F. Delbaen, and M. Kupper. Dynamic monetary risk measures for bounded discrete-time processes. *Electronic Journal of Probability*, 11(3):57–106, 2006.
- K. Clement-Nyns, E. Haesen, and J. Driesen. The impact of charging plug-in hybrid electric vehicles on a residential distribution grid. *IEEE Transactions on Power Systems*, 25(1):371–380, 2010.
- H. Dette, N. Neumeyer, and K. F. Pilz. A simple nonparametric estimator of a strictly monotone regression function. *Bernoulli*, 12(3):469–490, 2006.
- U. Eberle and R. von Helmolt. Sustainable transportation based on electric vehicle concepts: a brief overview. *Energy & Environmental Science*, 3(6):689, 2010.
- L. Gan, U. Topcu, and S. Low. Optimal decentralized protocol for electric vehicle charging. *IEEE Transactions on Power Systems*, 28(2):940–951, 2013.
- D. P. Heyman and M. J. Sobel. *Stochastic Models in Operations Research: Stochastic Optimization*, volume 2. Courier Corporation, 2003.
- J. E. Hogan, K. N. Lemon, and B. Libai. What is the true value of a lost customer? *Journal of Service Research*, 5(3):196–208, 2003.
- L. J. Hong and G. Liu. Simulating sensitivities of conditional value at risk. *Management Science*, 55(2):281–293, 2009.
- E. B. Iversen, J. M. Morales, and H. Madsen. Optimal charging of an electric vehicle using a Markov decision process. *Applied Energy*, 123:1–12, 2014.
- M. M. Karbasioun, I. Lambadaris, G. Shaikhet, and E. Kranakis. Optimal charging strategies for electrical vehicles under real time pricing. In *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 746–751, 2014.
- I. Koutsopoulos, V. Hatzi, and L. Tassiulas. Optimal energy storage control policies for the smart power grid. *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 475–480, 2011.
- V. Kozmík and D. P. Morton. Evaluating policies in risk-averse multi-stage stochastic programming. *Mathematical Programming*, 2014.
- R. Li, Q. Wu, and S. S. Oren. Distribution locational marginal pricing for optimal electric vehicle charging management. *IEEE Transactions on Power Systems*, 29(1):203–211, 2014.
- A. Megretski. SPOT: Systems polynomial optimization tools, 2013.
- H. Mukerjee. Monotone nonparametric regression. *The Annals of Statistics*, 16(2):741–750, 1988.

- P. A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming, Series B*, 96(2):293–320, 2003.
- G. C. Pflug. Some remarks on the value-at-risk and conditional value-at-risk. In *Probabilistic constrained optimization*, pages 272–281. Springer US, 2000.
- A. B. Philpott and V. L. de Matos. Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion. *European Journal of Operational Research*, 218(2):470–483, 2012.
- A. B. Philpott, V. L. de Matos, and E. Finardi. On solving multistage stochastic programs with coherent risk measures. *Operations Research*, 61(4):957–970, 2013.
- M. L. Puterman. *Markov decision processes: Discrete stochastic dynamic programming*. Wiley, New York, 2014.
- R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–41, 2000.
- C. Roe, E. Farantatos, J. Meisel, A. P. Meliopoulos, and T. Overbye. Power system level impacts of PHEVs. *2009 42nd Hawaii International Conference on System Sciences*, pages 1–10, 2009.
- N. Rotering and M. Ilic. Optimal charge control of plug-in hybrid electric vehicles in deregulated electricity markets. *IEEE Transactions on Power Systems*, 26(3):1021–1029, 2011.
- D. Roy. The discrete normal distribution. *Communications in Statistics – Theory and Methods*, 32(10):1871–1883, 2003.
- B. Rudloff, A. Street, and D. M. Valladão. Time consistency and risk averse dynamic decision models: definition, interpretation and practical consequences. *European Journal of Operational Research*, 234(3):743–750, 2014.
- A. Ruszczyński. Risk-averse dynamic programming for Markov decision processes. *Mathematical Programming*, 125(2):235–261, 2010.
- A. Schroeder and T. Traber. The economics of fast charging infrastructure for electric vehicles. *Energy Policy*, 43:136–144, 2012.
- N. Secomandi. Optimal commodity trading with a capacitated storage asset. *Management Science*, 56(3):449–467, 2010.
- M. Shaked and J. G. Shanthikumar. *Stochastic Orders*. Springer Science & Business Media., 2007.



- A. Shapiro, W. Tekaya, J. P. da Costa, and M. P. Soares. Risk neutral and risk averse stochastic dual dynamic programming method. *European Journal of Operational Research*, 224(2):375–391, 2013.
- E. Sortomme, M. M. Hindi, S. D. James MacPherson, and S. S. Venkata. Coordinated charging of plug-in hybrid electric vehicles to minimize distribution system losses. *IEEE Transactions on Smart Grid*, 2(1):186–193, 2011.
- N. Stockey and R. E. Lucas, Jr. *Recursive methods in economic dynamics*. Harvard University Press, 1989.
- O. Sundstrom and C. Binding. Optimization methods to plan the charging of electric vehicle fleets. *Proceedings of the International Conference on Control Communication and Power Engineering*, pages 28–29, 2010.
- U.S. Department of Energy. Plug-in electric vehicle handbook for public charging station hosts. Technical report, 2012.
- U.S. Department of Energy. Electric vehicle public charging – Time vs. energy. Technical report, 2013.
- Qi Wei, Xu Zhiwei, Z J M Shen, Hu Zechun, and Song Yonghua. Hierarchical coordinated control of plug-in electric vehicles charging in multifamily dwellings. *IEEE Trans. Smart Grid*, 5(3):1465–1474, 2014.
- M. D. Wytock, K. Adelberger, W. Wolf-Dietrich, and O. Irzak. Mediator device for smart electric vehicle charging. *US Patent Application No. 14/928,170*, 2016.
- T. Zhang, W. Chen, Z. Han, and Z. Cao. Charging scheduling of electric vehicles with local renewable energy under uncertain electric vehicle arrival and grid power price. *IEEE Transactions on Vehicular Technology*, 63(6):2600–2612, 2014.
- Y. H. Zhou, A. Scheller-Wolf, N. Secomandi, and S. Smith. Electricity trading and negative prices: Storage vs. disposal. *Management Science*, 2015.